SYSTEM AND METHOD FOR DISTURBING UNDESIRABLE TRANSMISSION OR RECEPTION OF ELECTRONIC MESSAGE

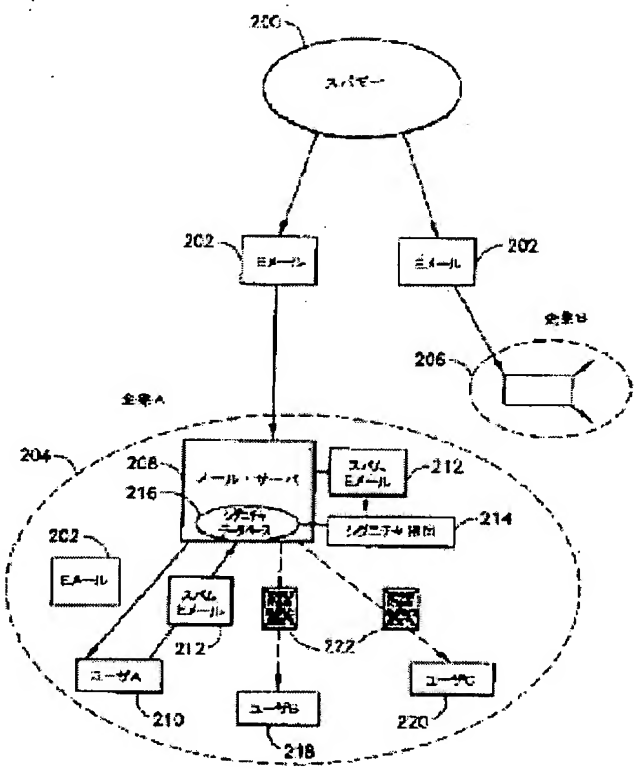
Patent number: JP2000353133
Publication date: 2000-12-19
Inventor: KEPHART JEFFREY OWEN
Applicant: IBM
Classification:
- international: ***G06Q10/00;*** *H04L12/58;* **G06Q10/00;** H04L12/58; (IPC1-7): G06F13/00; G06F15/00; H04L12/54; H04L12/58
- european: G06Q10/00F2
Application number: JP20000105418 20000406
Priority number(s): US19990289023 19990409

Also published as:
US6732149 (B1)
GB2350747 (A)

Report a data error here

Abstract of **JP2000353133**

PROBLEM TO BE SOLVED: To detect and process instances of undesirable mail of all types by scanning multiple inbound messages and confirming whether or not there is a specific electronic message, and taking a proper action in response to the scanning step. SOLUTION: A user 210 while reading a message and adding a label of spam 212 by regarding it as undesirable one scans its messages to confirm whether or not there are multiple undesirable messages which are already known. Messages of a user 218 are scanned several minutes later to confirm whether or not there are multiple undesirable messages 222. Further, a user 220 confirm 4th undesirable messages 222 30 minutes later. Thus, proper actions are taken in response to scanning steps. Consequently, instances of undesirable mail of all types can be detected.



D-2/9

P3014

(19)日本国特許庁（ＪＰ）　(12)**公開特許公報**（Ａ）　(11)特許出願公開番号
特開2000－353133
（P2000－353133A）
(43)公開日　平成12年12月19日(2000.12.19)

(51)Int.Cl.⁷		識別記号	FI		テーマコード(参考)
Ｇ０６Ｆ	13/00	３５１	Ｇ０６Ｆ	13/00	３５１Ｇ
	15/00	３１０		15/00	３１０Ｅ
Ｈ０４Ｌ	12/54		Ｈ０４Ｌ	11/20	１０１Ｂ
	12/58				

審査請求　有　　請求項の数45　ＯＬ　（全 24 頁）

(21)出願番号　特願2000－105418(P2000－105418)

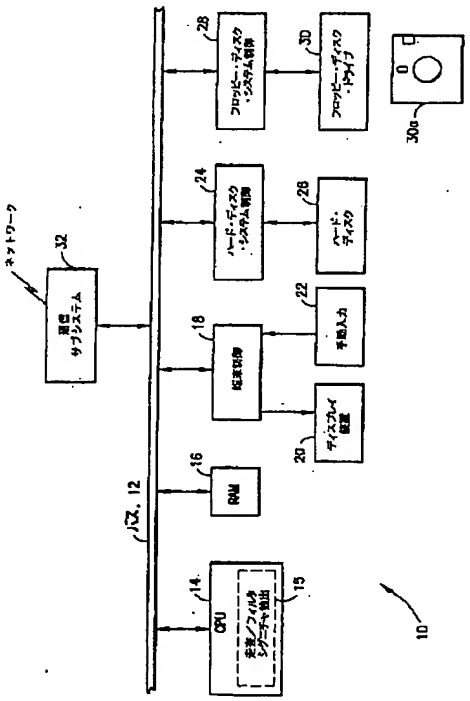
(22)出願日　平成12年４月６日(2000.4.6)

(31)優先権主張番号　０９／２８９０２３
(32)優先日　平成11年４月９日(1999.4.9)
(33)優先権主張国　米国（ＵＳ）

(71)出願人　390009531
インターナショナル・ビジネス・マシーンズ・コーポレーション
ＩＮＴＥＲＮＡＴＩＯＮＡＬ　ＢＵＳＩＮＥＳＳ　ＭＡＳＣＨＩＮＥＳ　ＣＯＲＰＯＲＡＴＩＯＮ
アメリカ合衆国10504、ニューヨーク州　アーモンク　（番地なし）
(72)発明者　ジェフリー・オーエン・ケファート
アメリカ合衆国10567　ニューヨーク州コートランド・マナー　ロバータ・ドライブ　14
(74)代理人　100086243
弁理士　坂口　博　（外１名）

(54)【発明の名称】　電子メッセージの望ましくない送信または受信を妨害するためのシステムおよび方法

(57)【要約】
【課題】　複数ユーザからなるネットワーク内で電子メッセージの望ましくない送信または受信を妨害するシステムおよび方法を提供すること。
【解決手段】　少なくとも１つの特定の電子メッセージの送信または受信が望ましくないものであることを判定するステップと、少なくとも１つの特定の電子メッセージまたはその変形の検出を可能にする検出データを自動的に抽出するステップと、少なくとも１人のユーザからの１つまたは複数のインバウンド・メッセージまたはアウトバウンド・メッセージあるいはその両方を走査して少なくとも１つの特定の電子メッセージまたはその変形が存在するかどうかを確認するステップと、走査ステップに応答して適切なアクションを実行するステップとを含む。



1

【特許請求の範囲】

【請求項１】複数ユーザからなるネットワーク内で電子メッセージの望ましくない送信または受信を妨害する方法において、

少なくとも１つの特定の電子メッセージの送信または受信が望ましくないものであることを判定するステップと、

前記少なくとも１つの特定の電子メッセージまたはその変形の検出を可能にする検出データを自動的に抽出するステップと、

少なくとも１人のユーザからの１つまたは複数のインバウンド・メッセージまたはアウトバウンド・メッセージあるいはその両方を走査して前記少なくとも１つの特定の電子メッセージまたはその変形が存在するかどうかを確認するステップと、

前記走査ステップに応答して適切なアクションを実行するステップとを含む方法。

【請求項２】抽出した前記検出データを記憶するステップをさらに含む、請求項１に記載の方法。

【請求項３】前記判定ステップが、前記少なくとも１つの特定の電子メッセージの増殖が望ましくないという通知を受信するステップを含む、請求項１に記載の方法。

【請求項４】前記受信ステップが、前記少なくとも１つの特定の電子メッセージを望ましくないものまたは機密のものとして識別する信号をアラート・ユーザから受信するステップを含む、請求項３に記載の方法。

【請求項５】前記少なくとも１つの特定の電子メッセージが前記アラート・ユーザの受信箱に受信される、請求項４に記載の方法。

【請求項６】前記受信ステップが、前記特定の電子メッセージに望ましくないものとしてのフラグを付けるべきであることを示すための識別子を前記アラート・ユーザに提供するステップを含む、請求項４に記載の方法。

【請求項７】前記提供ステップが、電子メッセージが望ましくないものであるという識別を援助するために総称検出器を提供するステップを含む、請求項６に記載の方法。

【請求項８】前記抽出ステップが、前記少なくとも１つの特定の電子メッセージからシグニチャ情報を抽出するステップを含む、請求項２に記載の方法。

【請求項９】前記記憶ステップが、前記走査ステップに応答して、前記少なくとも１つの特定の電子メッセージに関する情報を前記シグニチャ情報に追加するステップを含む、請求項８に記載の方法。

【請求項１０】前記抽出ステップが、前記少なくとも１つの特定の電子メッセージからシグニチャを抽出するステップを含む、請求項２に記載の方法。

【請求項１１】前記記憶ステップが、前記シグニチャを少なくとも１つのシグニチャ・データベースに記憶するステップを含む、請求項１０に記載の方法。

2

【請求項１２】前記シグニチャ・データベースが複数のシグニチャ・クラスタを含み、各クラスタが実質的に類似した電子メッセージに対応するデータを含む、請求項１１に記載の方法。

【請求項１３】前記抽出ステップおよび前記走査ステップが、複数ユーザからなる前記ネットワークの全域で同時かつ非同期に行われる、請求項２に記載の方法。

【請求項１４】前記走査ステップの前に、前記少なくとも１つの特定の電子メッセージが望ましくないものであることを確認するステップをさらに含む、請求項４に記載の方法。

【請求項１５】前記確認ステップが、総称検出技法により前記少なくとも１つの特定の電子メッセージが望ましくないものであることを確認するステップを含む、請求項１４に記載の方法。

【請求項１６】前記確認ステップが、前記少なくとも１つの特定の電子メッセージが望ましくないものであることを所定の限界数のユーザが通知することを必要とするステップを含む、請求項１４に記載の方法。

【請求項１７】前記走査ステップが、

メッセージ本文を抽出するステップと、

前記メッセージ本文を不変形式に変換するステップと、

前記不変形式を走査して前記検出データに対する正確な一致または接近した一致があるかどうかを確認するステップと、

各一致ごとに一致のレベルを決定するステップとを含む、請求項１に記載の方法。

【請求項１８】前記実行ステップが、前記少なくとも１つの特定の電子メッセージまたはその変形の存在を発見したときに適切なアクションを実行するステップを含む、請求項１に記載の方法。

【請求項１９】前記実行ステップが、前記少なくとも１つの特定の電子メッセージまたはその変形に望ましくないものまたは機密のものとしてのラベルを付けるステップを含む、請求項１８に記載の方法。

【請求項２０】前記実行ステップが、前記少なくとも１つの特定の電子メッセージまたはその変形を除去するステップを含む、請求項１８に記載の方法。

【請求項２１】前記実行ステップが、１つまたは複数のユーザ・プリファレンスに応答して、決定した一致のレベルごとに適切なアクションを実行するステップを含む、請求項１７に記載の方法。

【請求項２２】前記決定ステップが、

各一致ごとに最長領域一致を検出するステップと、

走査したメッセージのハッシュブロックと抽出した検出データのそれぞれのハッシュブロックの間のハッシュブロックの類似性を計算するステップと、

１つまたは複数のユーザ・プリファレンスを受信するステップと、

前記検出ステップ、前記計算ステップ、および前記受信

ステップに応答して、一致のレベルを決定するステップとを含む、請求項１７に記載の方法。

【請求項２３】複数ユーザからなるネットワーク内で電子メッセージの望ましくない送信または受信を妨害するための方法ステップを実行するためにマシンによって実行可能な命令からなるプログラムを具体的に実施する、マシンによって読取り可能なプログラム記憶装置において、前記方法が、

少なくとも１つの特定の電子メッセージの送信または受信が望ましくないものであることを判定するステップと、

前記少なくとも１つの特定の電子メッセージまたはその変形の検出を可能にする検出データを自動的に抽出するステップと、

少なくとも１人のユーザからの１つまたは複数のインバウンド・メッセージまたはアウトバウンド・メッセージあるいはその両方を走査して前記少なくとも１つの特定の電子メッセージまたはその変形が存在するかどうかを確認するステップと、

前記走査ステップに応答して適切なアクションを実行するステップとを含むプログラム記憶装置。

【請求項２４】複数ユーザからなるネットワーク内で電子メッセージの望ましくない送信または受信を妨害するためのシステムにおいて、

少なくとも１つの特定の電子メッセージの送信または受信が望ましくないものであることを判定する手段と、

前記少なくとも１つの特定の電子メッセージまたはその変形の検出を可能にする検出データを自動的に抽出する手段と、

少なくとも１人のユーザからの１つまたは複数のインバウンド・メッセージまたはアウトバウンド・メッセージあるいはその両方を走査して前記少なくとも１つの特定の電子メッセージまたはその変形が存在するかどうかを確認する手段と、

前記走査手段に応答して適切なアクションを実行する手段とを含むシステム。

【請求項２５】抽出した前記検出データを記憶する手段をさらに含む、請求項２４に記載のシステム。

【請求項２６】前記判定手段が、前記少なくとも１つの特定の電子メッセージの増殖が望ましくないという通知を受信する手段を含む、請求項２４に記載のシステム。

【請求項２７】前記受信手段が、前記少なくとも１つの特定の電子メッセージを望ましくないものまたは機密のものとして識別する信号をアラート・ユーザから受信する手段を含む、請求項２６に記載のシステム。

【請求項２８】前記少なくとも１つの特定の電子メッセージが前記アラート・ユーザの受信箱に受信される、請求項２７に記載のシステム。

【請求項２９】前記受信手段が、前記特定の電子メッセージに望ましくないものとしてのフラグを付けるべきであることを示すための識別子を前記アラート・ユーザに提供する手段を含む、請求項２７に記載のシステム。

【請求項３０】前記提供手段が、電子メッセージが望ましくないものであるという識別を援助するために総称検出器を提供する手段を含む、請求項２９に記載のシステム。

【請求項３１】前記抽出手段が、前記少なくとも１つの特定の電子メッセージからシグニチャ情報を抽出する手段を含む、請求項２５に記載のシステム。

【請求項３２】前記記憶手段が、前記走査手段に応答して、前記少なくとも１つの特定の電子メッセージに関する情報を前記シグニチャ情報に追加する手段を含む、請求項３１に記載のシステム。

【請求項３３】前記抽出手段が、前記少なくとも１つの特定の電子メッセージからシグニチャを抽出する手段を含む、請求項２５に記載のシステム。

【請求項３４】前記記憶手段が、前記シグニチャを少なくとも１つのシグニチャ・データベースに記憶する手段を含む、請求項３３に記載のシステム。

【請求項３５】前記シグニチャ・データベースが複数のシグニチャ・クラスタを含み、各クラスタが実質的に類似した電子メッセージに対応するデータを含む、請求項３４に記載のシステム。

【請求項３６】前記抽出手段および前記走査手段が、複数ユーザからなる前記ネットワークの全域で同時かつ非同期に処理する、請求項２５に記載のシステム。

【請求項３７】前記少なくとも１つの特定の電子メッセージが望ましくないものであることを確認する手段をさらに含む、請求項２５に記載のシステム。

【請求項３８】前記確認手段が、総称検出技法により前記少なくとも１つの特定の電子メッセージが望ましくないものであることを確認する手段を含む、請求項３７に記載のシステム。

【請求項３９】前記確認手段が、前記少なくとも１つの特定の電子メッセージが望ましくないものであることを所定の限界数のユーザが通知することを必要とする手段を含む、請求項３７に記載のシステム。

【請求項４０】前記走査手段が、

メッセージ本文を抽出する手段と、

前記メッセージ本文を不変形式に変換する手段と、

前記不変形式を走査して前記検出データに対する正確な一致または接近した一致があるかどうかを確認する手段と、

各一致ごとに一致のレベルを決定する手段とを含む、請求項２４に記載のシステム。

【請求項４１】前記実行手段が、前記少なくとも１つの特定の電子メッセージまたはその変形の存在を発見したときに適切なアクションを実行する手段を含む、請求項２４に記載のシステム。

【請求項４２】前記実行手段が、前記少なくとも１つの

5

特定の電子メッセージまたはその変形に望ましくないものまたは機密のものとしてのラベルを付ける手段を含む、請求項４１に記載のシステム。

【請求項４３】前記実行手段が、前記少なくとも１つの特定の電子メッセージまたはその変形を除去する手段を含む、請求項４１に記載のシステム。

【請求項４４】前記実行手段が、１つまたは複数のユーザ・プリファレンスに応答して、決定した一致のレベルごとに適切なアクションを実行する手段を含む、請求項４０に記載のシステム。

【請求項４５】前記決定手段が、

各一致ごとに最長領域一致を検出する手段と、

走査したメッセージのハッシュブロックと抽出した検出データのそれぞれのハッシュブロックの間のハッシュブロックの類似性を計算する手段と、

１つまたは複数のユーザ・プリファレンスを受信する手段と、

前記検出手段、前記計算手段、および前記受信手段に応答して、一致のレベルを決定する手段とを含む、請求項４０に記載のシステム。

【発明の詳細な説明】

【０００１】

【発明の属する技術分野】本発明は、一般に電子メールおよびその他のタイプの電子メッセージの送信および受信が可能なディジタル・データ・プロセッサおよび相互通信するディジタル・データ・プロセッサのネットワークに関する。特に、本発明は、「スパム」（spam）とも呼ばれる非送信請求商用Ｅメール（ＵＣＥ）など、望ましくない非送信請求電子メールを自動的に検出し処理するためのシステムおよび方法に関する。

【０００２】

【従来の技術】毎日、無数のインターネット・ユーザは、通常、電子メール（Ｅメール）の形でありがたくない電子メッセージを受信する。このようなメッセージの最もありふれた例は、一般に「スパム」と呼ばれる非送信請求商用Ｅメール（ＵＣＥ）である。ＵＣＥは、通常、特定の商品、サービス、またはウェブ・サイトを奨励するものであり、数千人または数百万人もの個人に無差別に送信され、その大多数はＵＣＥが迷惑なものまたは不快なものと認識している。ＵＣＥは、重大な問題として広く認められている。ＵＣＥに関する記事はほとんど毎日のようにＣＮＥＴなどの技術ニュース・サービスに登場する。Ｅメール・ユーザがＵＣＥに曝される機会を低減するために、いくつかの市販用製品およびシェアウェア製品が作成された。少なくとも１社の新設企業であるBright Light Technologiesは、ＵＣＥを検出してフィルタリングで除去するための技術を生産して販売するという唯一の目的のために設立された。法律による制限はいくつかの州で企図されており、実際に複数の州で最近実施された。

6

【０００３】その他の形式の望ましくないＥメールとしては、うわさ、でっちあげ、チェーン・レターなどがある。このような形式のＥメールのそれぞれは、複数ユーザからなるネットワーク内で非常に素早く増殖する可能性がある。うわさは、ユーザ全体に非常に勢いよく広まる可能性があり、その結果、時間の浪費や不必要な懸念を引き起こす可能性がある。でっちあげのコンピュータ・ウィルスとして最も上出来なものは、コンピュータ・ウィルスそのものに匹敵するほど長続きし、相当なパニックを引き起こす恐れがある。最後に、チェーン・レターの流布は、企業の方針または連邦法によって禁止するのに十分なほど重大な現象である。

【０００４】幾分異なるクラスのＥメールは機密Ｅメールであり、その送信または受信は望ましくない場合が多い。機密Ｅメールは、選択されたグループ外の人には転送されないようになっている。したがって、これらのメッセージの配布を制御することに関する懸念が発生する。

【０００５】ＵＣＥおよび電子的に伝えられるうわさ、でっちあげ、チェーン・レターに共通の特徴は、問題のメッセージの内容（およびその送信）が（単に面白くないどころか）望ましくないものであるという合意が広まりそうなことである。これは、このようなメッセージが電子的形式になっているということとともに、このＥメールを自動的に検出し、無害なものにしようと試みる様々な技術を企図することを可能にするものである。

【０００６】現在までのところ、ＵＣＥは、このような努力の排他的な焦点になっていた。既存のＵＣＥソリューションはいくつかの異なる形を取る。その一部は、既存のＥメール・パッケージ（たとえば、Eudoraメール・システムとともに機能するよう設計されたMailJail）またはＥメール・プロトコル（たとえば、Windows 95、Windows 98、またはWindows NTプラットフォーム上でＰＯＰ３プロトコルをサポートするいずれかのＥメール・パッケージのために機能するSpam Exterminator）とともに機能するよう設計されたソフトウェア・パッケージである。その他のソリューションは、広く使用されるメール・プロトコル（たとえば、指定のサイトからあるいは明示的に許されているもの以外のサイトからのメール・リレーをブロックするための機構を提供するSendMailメール転送プロトコルの最近のアップグレードであるSendMail v.8.8）に統合されている。もう１つのタイプのソリューションは、Ｅメール・フィルタリング・サービス、たとえば、junkproof.comによって提供されるサービスであり、これはＵＣＥを送信したユーザに罰金を科すものである。Bright Light Technologiesでは、ソフトウェア製品とサービスを併用するよう提案している。

【０００７】どのようにパッケージ化可能であっても、これらのソリューションの大部分は、認識と応答という２つの主なステップから構成される。認識ステップで

は、所与のＥメール・メッセージを検査して、それがスパムになりそうかどうかを判定する。認識ステップ中にそのメッセージがスパムになりそうであると見なした場合、何らかの応答を行う。典型的な応答としては、そのメッセージを自動的に削除すること、それがスパムである可能性があることにユーザの注意をひくようそれにラベルまたはフラグを付けること、優先順位がより低いメール・フォルダにそれを入れることなどがあり、おそらく、カストマイズ可能なメッセージを送信側に返送することと結合される。

【０００８】主な技術的難題は認識ステップにある。最も重要な難題のうちの２つは、偽陽性（合法的メールを偽ってスパムとして告発すること）および偽陰性（スパムをそれとして識別し損なうこと）の率をできるだけ低く維持することを含む。多種多様な市販用アプリケーションおよびフリーウェア・アプリケーションでは、一般的な問題に対処するため、以下の基本的なスパム検出戦略に基づく組合せまたは変形あるいはその両方を使用する。

【０００９】ドメインベースの検出

スパムを送信する人（「スパマー」（spammers））は、そこからスパムを送信するための特別なインターネット・アドレス・ドメインをセットアップする場合が多い。一般的な反スパム・ソリューションの１つは、「スパム」ドメインのブラックリストを維持し、このようなドメインから発信されたメールを拒否するか、送達しないか、または送信側に返すことである。スパムが新しい「スパム」ドメインから流出し始めると、そのドメインをブラックリストに追加することができる。

【００１０】たとえば、xmission.comでは、指定のサイトからのメールを送信側に返送するようにsendmail.cfの規則を変更した。そのテキスト・ファイル（http://spam.abuse.net/spam/tools/dropbad.txt）は、moneyworld.com、cyberpromo.com、bulk-e-mail.com、bigprofits.comなどを含む、スパマーによる使用のためにのみセットアップされたことが分かっているいくつかのドメインをリストしている。http://www.webeasy.com:8080/spam_download_tableには、１０００箇所以上のサイトがこのようにブラックリスト化されている。SendMailの最近のバージョン（バージョン８．８以降）はこのようなリストを使用を容易にするよう変更されており、これはスパムとの戦いにおいて重要な進展と見なされている。

【００１１】しかし、無差別に使用されると、この手法の結果、偽陽性率および偽陰性率が高くなる可能性がある。たとえば、スパマーがaol.comドメインからスパムを送信する場合、aol.comはブラックリストに追加することもできる。その結果、このドメインから合法的にメールを送信する無数の個人はそのメールがブロックされる恐れがある。言い換えれば、偽陽性率は容認できないほど高くなる恐れがある。これに反して、スパマーは禁止ドメインから新たに作成した非禁止ドメインまたは多くの合法的ユーザが使用するドメインへ素早く切り替えることができ、したがって、多くの偽陰性が発生する。

【００１２】ヘッダベースの検出

スパムの顕著な特徴は、極めて多数の受信側に送信されることである。これを示す表示であり、あるメッセージがスパムになりそうであるという証拠と見なすことができるものがメール・メッセージのヘッダ内に存在する場合が多い。たとえば、長い受信側のリストは通常、集合名のより小さいセットに送信することによって処理されるので、そのユーザの明示的なＥメール・アドレスはTo:フィールドに現れない。

【００１３】Ross Rader of Internet Direct（Idirect）は、Eudora Light、Microsoft Mail、Pegasusを含む、普及している多様なＥメール・プログラムのためにスパムのこの特徴に基づいて単純な規則をセットアップするための説明書を発表した。メール・メッセージ・ヘッダがその規則と一致する場合、そのメールはユーザの受信箱から自動的に除去され、特別なフォルダ内に置かれるが、そのフォルダではそれを後で調べるかまたは検査なしで容易に削除することができる。

【００１４】しかし、この方法のユーザがこれらの検出規則の個別設定に多くの努力を注がない限り、偽陽性率が非常に高くなる可能性があり、したがって、合法的Ｅメールの大部分はスパムと分類されることになる。

【００１５】テキストベースのキーワード検出

スパムは、通常、製品を販売し、ポルノのウェブ・サイトを訪れることを唱え、ネズミ講式販売またはその他の金銭的詐欺への賛助を得ようと積極的に試みるという点で普通のＥメールとは区別される。したがって、「MAKE MONEY FAST」というテキスト・フラグメントを含む１通のメールは、「During my meeting with you last Tuesday」で始まるメールよりスパムである可能性が高い。

【００１６】一部の反スパム方法は、各Ｅメールの本文を走査して、スパムには見られるが他のＥメールには見られない傾向のあるキーワードまたはキーフレーズを検出する。キーワード・リストおよびキーフレーズ・リストはカストマイズ可能である。この方法は、上記のドメインベース検出技法およびヘッダベース検出技法と併用される場合が多い。この技法の例としては、procmail、Spam Exterminator、SPAM Attack Pro!とともに機能するjunkfilter（http://www.pobox.com/gsutter/junkmail）がある。

【００１７】この場合も、普通のＥメール・メッセージが禁止キーワードまたはキーフレーズを含むときに偽陽性が発生する可能性がある。この手法は、偽陰性も被りがちである。というのは、禁止キーフレーズのリストは、スパムの新しいインスタンスの到来に遅れないようにするために、毎日数回更新しなければならない恐れが

9

あり、これは反スパム・ベンダにとって技術的に難しいと同時にユーザにとっても不快であるからである。

【００１８】テキストベースのマシン分類

Spam Be Gone!はEudoraとともに機能するフリーウェア製品である。これは、スパムおよび非スパムＥメールの例を記録し、それぞれのインスタンスに対する各着信Ｅメールの類似性を測定し、スパムまたは非スパムとしてのＥメールの分類に達するようその類似性のスコアを結合するインスタンスベースの分類プログラムである。この分類プログラムは、各個別ユーザごとに自動的にトレーニングされる。ユーザが分類プログラムを開発するには、通常、数週間から数カ月はかかる。

【００１９】十分な量のトレーニングの後、この手法の偽陽性率および偽陰性率は他の技法より低くなると言われている。ある事例（http://www.internz.com/SpamBeGone/stats.html）では、複数のユーザに関する平均が提供されないので、これがパフォーマンスに関する上限になると想定することができるが、偽陰性率は１ないし２カ月間のトレーニング後に１０分の数パーセントより低くなり、偽陽性率は１カ月後に２０％で、２カ月後に５％であった。したがって、最良の場合でも、スパムとしてラベルが付けられた２０のメッセージのうちの１つは実際には合法的なものになる。これは、反スパム・ソフトウェアが自動的にメールを削除するかまたはそれを送信側に返すなどの強い態度で応答する場合には特に容認できないものになる可能性もある。

【００２０】上記のＵＣＥ検出方法はいずれも、スパムにとって総称的であるが普通の非スパムＥメールではあまり一般的ではない機能を使用するという意味で「総称的」なものである。これは、通常、ホスト・プログラムを走査して特定のウィルスを示す特別な「シグニチャ」バイト・パターンがあるかどうかを確認することにより特定の既知のコンピュータ・ウィルスを検出するためにウィルス対策プログラムによって一般に使用される「特定の」検出技法とは対照的なものである。総称認識技法は、以前は未知であった新しいスパムを捕らえることができるので、魅力的なものである。しかし、上記で示したように、その欠点は、容認できないほど高い偽陽性率を生じ、場合によっては容認できないほど高い偽陰性率も生じる傾向があることである。特定の検出技法は、通常、より小さい偽陽性率および偽陰性率を有するが、総称技法を実行するより頻繁な更新が必要になる。

【００２１】総称検出技法は、うわさ、でっちあげ、チェーン・レターまたは機密Ｅメールなど、他のタイプの望ましくないＥメールを認識する際にあまり役に立ちそうもない。送信側のドメインまたはメール・ヘッダのその他の態様に基づく認識は、まったく機能しそうもない。メッセージ本文内に存在するキーワードまたはキーフレーズを基礎とするでっちあげおよびチェーン・レターの総称認識は可能であるかもしれないが、内容の範囲

10

がより広くなりそうなので、スパムの場合より難しくなりそうである。また、テキストを基礎とする機密Ｅメールの総称認識はほとんど確実に不可能である。というのは、どのようなマシン・アルゴリズムでも認識可能なやり方で機密テキストと非機密テキストを区別するものが何もないからである。

【００２２】Bright Light Technologiesは他の反スパム製品／サービスを奨励している。Bright Lightでは、インターネット全体で複数のＥメール・アドレス（または「プローブ」）を使用しているが、これは合法的宛先ではないので、理論上は望ましくないメッセージのみを受信するものである。受信したメッセージは、２４時間体制のオペレーション・センタに配置されたオペレータによって読まれる。これらのオペレータは、メッセージを評価し、ユーザ・グループに対応するメール・サーバ内のスパムブロック機能を制御する規則を更新する。

【００２３】ＵＣＥ検出および応答のこの方法は、総称検出ではなく特定の検出を使用するので、本質的には偽陽性および偽陰性に対してあまり脆弱ではないが、いくつかの欠点がある。そのうちの多くは、サービスを維持するために必要な大量の手動労力に由来するものである。Bright Lightのオペレーション・センタは、スパムがあるかどうかＥメールのストリームをモニタし、スパムの特定の事例を示す適切なインジケータと思われるキーワードおよびキーフレーズを手動で抽出し、これらのキーワードまたはキーフレーズをデータベースに記憶する専門家を使用しなければならない。いずれかの企業がこのような１組の専門家を独力でサポートすることはほとんど法外に高くなりそうなので、このようにして自社を保護することを希望する企業は、Bright Lightのオペレーション・センタによる連続し途切れないサービスに完全に依存することになると思われる。少なくとも一部の企業は、外部組織からより自由であることと、単一組織によって達成されそうなもの以上のカストマイズを見込んでいるソリューションを選ぶかもしれない。この問題の最重要点は、Bright Lightの方法では互いに独立しているべき２つのタスク、すなわち、望ましくないものとしてメッセージにラベルを付けることと、望ましくないメッセージからシグニチャを抽出することが結合されることである。手動入力の要件を望ましくないメッセージにラベルを付ける際の要件に削減することが可能である場合、これにより、望ましくないメッセージの協調的決定のローカライズが可能になると思われる。さらに、Bright Lightは、キーワードまたはフレーズに基づいて可能な一致を特定のメッセージ全体（またはその大部分）への正確な一致または近似の一致によってより厳重にテストすることができるような補助データを専門家が抽出する際のプロセスについて説明していない。したがって、その特定のソリューションは、個別ユーザがメッセージ一致に関するより厳重な条件を指定する機会を持

11

つと思われるものより偽陽性に対してより脆弱になりそうである。

【００２４】もう１つの欠点は、でっちあげ、チェーン・レター、不適当に転送された機密メッセージを含むより広範囲のクラスの望ましくないメッセージとは対照的に、Bright Lightのソリューションは特にＵＣＥに向けられていることである。ひとまとめにして考えると、プローブ・アカウントはすべてのＵＣＥのうちの妥当な断片を受信することができるが、それがチェーン・レターおよびうわさを引き寄せることは不明である。

【００２５】

【発明が解決しようとする課題】したがって、本発明の一目的は、非常に低い偽陽性率および偽陰性率で、すべてのタイプの望ましくないメールのインスタンスを検出し処理するための自動的かつ非総称的な手順を提供することにある。

【００２６】本発明の他の目的は、スタッフ配置を伴わず、むしろ積極的にＵＣＥを識別するためにユーザ自身を利用する安価なソリューションを提供することにある。

【００２７】本発明のさらに他の目的は、機密Ｅメール・メッセージの望ましくない送信または受信あるいはその両方を防止するためのシステムおよび方法を提供することにある。

【００２８】

【課題を解決するための手段】本発明は、低い偽陰性率および非常に低い偽陽性率でＵＣＥおよび他の形式の望ましくないＥメールを正確に検出し処理するための自動手順を提供する。既存の総称検出方法とは対照的に、本発明は、特定の検出技法を使用して望ましくないメッセージを認識する。言い換えれば、本発明のシステムは、望ましくないメッセージの特定のインスタンスに対するその正確な一致または接近した一致を基礎として望ましくないメッセージを効率よく検出する。Bright Lightによって使用される特定の技法とは対照的に、特定の望ましくないメッセージを識別するために使用する文字ストリングは完全に自動的に導出され、エンド・ユーザが様々なレベルの応答を開始するのに必要な一致度を調整できるようにする補助データで補足される。もう１つの対照的な点は、シグニチャ・データの自動導出によりフレキシビリティが増すことである。というのは、必要な唯一の手動入力は望ましくないものとして特定のメッセージにラベルを付けることであるからである。このため、普通のユーザは、望ましくないメッセージを定義するために協調的に機能することができ、専門家が手動でラベルを付け、望ましくないメッセージからシグニチャを抽出しなければならない外部の集中オペレーション・センタに対する依存状態から解放される。これにより、でっちあげおよびチェーン・レターの権威は、まったく異なる種類の専門知識を必要とする恐れがある、シグニチャ

12

の抽出という負担をさらに課すことなしに、それを含むメッセージを識別することができる。もう１つの対照的な点は、抽出したシグニチャ・データによりユーザが、シグニチャの一致からメッセージ全体の一語一句の一致に至る、所与のレベルの一致を構成するものについて独立したフレキシブルな定義を定義できることである。

【００２９】本発明の方法は、第１の（「アラート」）ユーザが望ましくないメールの所与のインスタンスを受信したときに、そのメッセージに望ましくないものとしてラベルを付けることと、そのメッセージに関するシグニチャを抽出することと、シグニチャ・データベースにそのシグニチャを追加することと、第２の（できる限り同じものを含む）ユーザのメッセージを定期的に走査してデータベース内のシグニチャに存在するかどうかを確認することと、第２のユーザのメッセージのうち、シグニチャを含むものを望ましくないものとして識別することと、このようにラベルが付けられたメッセージに対して適切に応答することとを含む。

【００３０】具体的には、複数ユーザからなるネットワーク内で電子メッセージの望ましくない送信または受信を妨害する方法は、少なくとも１つの特定の電子メッセージの送信または受信が望ましくないものであることを判定するステップと、少なくとも１つの特定の電子メッセージまたはその変形の検出を可能にする検出データを自動的に抽出するステップと、少なくとも１人のユーザからの１つまたは複数のインバウンド・メッセージまたはアウトバウンド・メッセージあるいはその両方を走査して少なくとも１つの特定の電子メッセージまたはその変形が存在するかどうかを確認するステップと、走査ステップに応答して適切なアクションを実行するステップとを含む。好ましいことに、この方法は、抽出した検出データを記憶するステップをさらに含む。

【００３１】好ましいことに、判定ステップは、少なくとも１つの特定の電子メッセージの増殖が望ましくないという通知を受信するステップを含む。この受信ステップは、好ましいことに、少なくとも１つの特定の電子メッセージを望ましくないものまたは機密のものとして識別する信号をアラート・ユーザから受信するステップを含む。少なくとも１つの特定の電子メッセージはアラート・ユーザの受信箱に受信することができる。この受信ステップは、好ましいことに、特定の電子メッセージに望ましくないものとしてのフラグを付けるべきであることを示すための識別子をアラート・ユーザに提供するステップを含む。この提供ステップが電子メッセージが望ましくないものであるという識別を援助するために総称検出器を提供するステップを含むことは、好ましいことである。

【００３２】本発明の抽出ステップは、好ましいことに、少なくとも１つの特定の電子メッセージからシグニチャ情報を抽出するステップを含む。記憶ステップは、

好ましいことに、走査ステップに応答して、少なくとも１つの特定の電子メッセージに関する情報をシグニチャ情報に追加するステップを含む。このシグニチャ情報は、好ましいことに、少なくとも１つの特定の電子メッセージからのシグニチャを含む。記憶ステップは、そのシグニチャを少なくとも１つのシグニチャ・データベースに記憶するステップを含むことができる。このシグニチャ・データベースは、好ましいことに、複数のシグニチャ・クラスタを含み、各クラスタは実質的に類似した電子メッセージに対応するデータを含む。シグニチャ・クラスタのそれぞれは、好ましいことに、走査情報を有する文字シーケンス・コンポーネントと、特定のシグニチャ変形に関する識別情報を有する原型コンポーネントとを含む。走査情報は、好ましいことに、特定の電子メッセージに関するサーチ文字シーケンスと、そのクラスタ内に表されるすべての電子メッセージに関する拡張文字シーケンス情報とを含み、識別情報は、特定のシグニチャ変形に関連する電子メッセージのフル・テキスト記憶コピーを指すポインタと、電子メッセージのハッシュブロックと、電子メッセージのコピーが受信され、その増殖が望ましくないものとしてアラート・ユーザによって報告された特定のインスタンスに対応するアラート・データとを含む。

【００３３】本発明の抽出ステップおよび走査ステップは、複数ユーザからなるネットワークの全域で同時かつ非同期に行うことができる。

【００３４】本発明の方法は、走査ステップの前に、少なくとも１つの特定の電子メッセージが望ましくないものであることを確認するステップをさらに含むことができる。この確認ステップは、好ましいことに、総称検出技法により少なくとも１つの特定の電子メッセージが望ましくないものであることを確認するステップを含む。請求項１４に記載の方法によれば、この確認ステップは、少なくとも１つの特定の電子メッセージが望ましくないものであることを所定の限界数のユーザが通知することを必要とするステップを含む。

【００３５】抽出ステップは、好ましいことに、特定の電子メッセージを走査して少なくとも１つのシグニチャ・データベース内のシグニチャがあるかどうかを確認するステップと、走査ステップで一致シグニチャを検出したことに応答して、その一致シグニチャを一致クラスタ内の各メッセージ変形と比較するステップとを含む。この比較ステップは、好ましいことに、特定の電子メッセージに関するハッシュブロックを計算するステップと、計算したハッシュブロックを各原型コンポーネントの識別情報内の変形ハッシュブロックと比較するステップとを含む。本発明の方法が、正確な変形ハッシュブロック一致が検出された場合に、そのポインタを使用して変形一致のフル・テキスト記憶コピーを検索するステップと、変形一致のフル・テキスト記憶コピーと特定の電子

メッセージのフル・テキストがその特定の電子メッセージを変形のインスタンスと見なすのに十分なほど類似していると見なされた場合に、特定の電子メッセージからアラート・データを抽出し、それを変形一致に関するアラート・データに追加するステップと、正確な変形ハッシュブロック一致が検出されないかまたは特定の電子メッセージのフル・テキストがそのデータベース内の変形のいずれかと十分に類似していないと判断された場合に、その特定の電子メッセージがいずれかの既存のクラスタと十分に類似しているかどうかを判定するステップと、その特定の電子メッセージがある既存のクラスタと十分に類似している場合に、特定の電子メッセージに関連する新しい識別情報を計算するステップと、その特定の電子メッセージが既存のクラスタと十分に類似していると判定されない場合に、その特定の電子メッセージに関する新しいクラスタを作成するステップとをさらに含むことは、好ましいことである。判定ステップは、好ましいことに、各クラスタの拡張文字シーケンス情報に示されたその特定の電子メッセージの領域のチェックサムを計算するステップと、計算したチェックサムを各クラスタの拡張文字シーケンス情報内の記憶チェックサムと比較するステップとを含む。この方法は、好ましいことに、シグニチャ一致がまったく検出されない場合に、その特定の電子メッセージに関する新しいクラスタを作成するステップをさらに含む。拡張文字シーケンス情報は、好ましいことに、開始オフセット・フィールドと、領域長フィールドと、ＣＲＣフィールドとを含み、この方法は、各クラスタごとに、最長領域長を有する一致領域を決定するステップと、すべてのクラスタのうちで最長領域長が少なくとも指定の限界長と等しい場合に、最長領域長クラスタを特定の電子メッセージ原型が追加される原型クラスタとして識別するステップとをさらに含む。最後に、本発明の方法は、好ましいことに、識別したクラスタの走査情報を再計算するステップを含む。アラート・データは、好ましいことに、コピーが本来受信された時刻を有する受信時刻フィールドを含み、この方法は、各シグニチャ・クラスタのすべての変形の受信時刻フィールドを現在時と定期的に比較するステップと、受信時刻フィールドのいずれも所定の日時より最近のものではないシグニチャ・クラスタを除去するステップとをさらに含む。

【００３６】走査ステップは、好ましいことに、メッセージ本文を抽出するステップと、メッセージ本文を不変形式に変換するステップと、不変形式を走査して検出データに対する正確な一致または接近した一致があるかどうかを確認するステップと、各一致ごとに一致のレベルを決定するステップとを含む。

【００３７】実行ステップは、好ましいことに、少なくとも１つの特定の電子メッセージまたはその変形の存在を発見したときに適切なアクションを実行するステップ

を含む。その実行ステップは、少なくとも１つの特定の電子メッセージまたはその変形に望ましくないものまたは機密のものとしてのラベルを付けるステップを含むことができる。また、この実行ステップは、少なくとも１つの特定の電子メッセージまたはその変形を除去するステップも含む。

【００３８】実行ステップは、好ましいことに、１つまたは複数のユーザ・プリファレンスに応答して、決定した一致のレベルごとに適切なアクションを実行するステップを含み、その決定ステップは、好ましいことに、各一致ごとに最長領域一致を検出するステップと、走査したメッセージのハッシュブロックと抽出した検出データのそれぞれのハッシュブロックの間のハッシュブロックの類似性を計算するステップと、１つまたは複数のユーザ・プリファレンスを受信するステップと、検出ステップ、計算ステップ、および受信ステップに応答して、一致のレベルを決定するステップとを含む。

【００３９】本発明は、複数ユーザからなるネットワーク内で電子メッセージの望ましくない送信または受信を妨害するための方法ステップを実行するためにマシンによって実行可能な命令からなるプログラムを具体的に実施する、マシンによって読取り可能なプログラム記憶装置も含み、その方法は、少なくとも１つの特定の電子メッセージの送信または受信が望ましくないものであることを判定するステップと、少なくとも１つの特定の電子メッセージまたはその変形の検出を可能にする検出データを自動的に抽出するステップと、少なくとも１人のユーザからの１つまたは複数のインバウンド・メッセージまたはアウトバウンド・メッセージあるいはその両方を走査して少なくとも１つの特定の電子メッセージまたはその変形が存在するかどうかを確認するステップと、走査ステップに応答して適切なアクションを実行するステップとを含む。

【００４０】最後に、本発明は、複数ユーザからなるネットワーク内で電子メッセージの望ましくない送信または受信を妨害するためのシステムであって、少なくとも１つの特定の電子メッセージの送信または受信が望ましくないものであることを判定する手段と、少なくとも１つの特定の電子メッセージまたはその変形の検出を可能にする検出データを自動的に抽出する手段と、少なくとも１人のユーザからの１つまたは複数のインバウンド・メッセージまたはアウトバウンド・メッセージあるいはその両方を走査して少なくとも１つの特定の電子メッセージまたはその変形が存在するかどうかを確認する手段と、走査手段に応答して適切なアクションを実行する手段とを含むシステムも含む。その他の点では、このシステムの好ましい実施の形態は本発明の方法の好ましい実施の形態と一致する。

【００４１】

【発明の実施の形態】図１は、本発明の教示を実施するために適当なシステム１０のブロック図である。バス１２は、中央演算処理装置（ＣＰＵ）１４と複数の他のシステム・バス・ユニットの間でアドレス、データ、制御を搬送するための複数の信号線からなる。ランダム・アクセス・メモリ（ＲＡＭ）１６は、システム・バス１２に結合され、プログラム命令記憶域および作業メモリをＣＰＵ１４に提供する。シグニチャ抽出モジュールおよび走査／フィルタ・モジュール１５は、その方法については以下に説明するが、ＣＰＵ１４あるいは別々のＣＰＵ上で動作することができる。端末制御サブシステム１８は、システム・バス１２に結合され、ディスプレイ装置２０、通常はＣＲＴまたはＬＣＤモニタに出力を供給し、キーボードまたはポインティング・デバイスなどの手動入力装置２２からの入力を受け取る。ハード・ディスク制御サブシステム２４は、回転固定ディスクまたはハード・ディスク２６をシステム・バス１２に両方向に結合する。制御装置２４およびハード・ディスク２６は、ＣＰＵ命令およびデータのための大容量記憶域を提供する。フロッピー・ディスク・ドライブ３０とともに、フロッピー・ディスケット３０ａからシステム・メモリへのコンピュータ・ファイルの転送の際に入力手段として有用なフロッピー・ディスク制御サブシステム２８は、フロッピー・ドライブ３０をシステム・バス１２に両方向に結合する。最後に、通信サブシステム３２は、システム・バス１２に結合され、インターネットなどのネットワークへのリンクを提供する。

【００４２】図１に示す構成要素は、パーソナル・コンピュータ、ポータブル・コンピュータ、ワークステーション、ミニコンピュータ、またはスーパーコンピュータ内で実施することができる。このため、バス１２の構造またはそのバスに結合されるＣＰＵ１４の数など、データ処理システム１０の物理的な実施の形態の詳細は、本発明の動作にとって重大なものではなく、以下ではこれ以上詳細には説明しない。

【００４３】大まかに言えば、本発明の方法は２つの段階を含む。第１に、シグニチャ抽出段階では、現在、システムによってそれとして認識されていないような望ましくない（または機密の）メッセージは、おそらく自動手順によって援助される第１のアラート・ユーザによって望ましくない（または機密の）ものとしてラベルが付けられ、所与のシグニチャ・データは、そのメッセージから自動的に抽出され、ユーザ全体に配布される１つまたは複数のデータベース内に置かれる。第２に、シグニチャ走査段階では、少なくとも１人のユーザのメッセージ・セット（できる限り第１のアラート・ユーザのセットを含む）は、実質的に類似したメッセージのインスタンスを検出しようとして、抽出したシグニチャ・データを使用して走査され、このようなメッセージが検出されると適切なアクションが実行される。

【００４４】図２は、具体的にスパムに対処する本発明

の一実施の形態が適用されるコンピュータ・システム環境を示している。スパマー２００は、スパム２０２を企業Ａ２０４および企業Ｂ２０６に送信する。実際には、スパム２０２は、多くの様々な企業に送信されるはずである。企業Ａ２０４が本発明を使用すると想定すると、スパム２０２は、１人または複数のユーザがアカウントを維持しているメール・サーバ２０８で受信される可能性がある。ユーザＡ２１０が自分のメールにアクセスすると想定すると、スパム２０２はその着信メールのリストで検出される。ユーザＡ２１０がスパム２０２をそれとして識別したことに応答して、識別されたスパム２１２はそれとしてラベルが付けられ、本発明のシグニチャ抽出段階が開始される。

【００４５】本発明のシグニチャ抽出段階では、識別されたスパム２１２はメール・サーバ２０８によってシグニチャ抽出エンジン２１４に転送される可能性がある。シグニチャ抽出エンジン２１４によって抽出された後、識別されたスパム２１２のシグニチャはメール・サーバ２０８に返され、シグニチャ・データベース２１６に記憶される。本発明のシグニチャ走査段階では、ユーザＢ２１８およびユーザＣ２２０の着信（または発信）メッセージは、シグニチャ・データベース２１６内の抽出シグニチャ・データを使用して走査される。この場合、実質的に類似したメッセージ２２２のインスタンスは、ユーザのためにフラグが付けられるか、その受信箱から除去されるか、または送信が防止される。

【００４６】この２通りの段階は、ユーザ全体にわたって同時かつ非同期に機能することができる。たとえば、ユーザＡは、メッセージ３を読み取って、それに望ましくないものとしてのラベルを付けながら、自分のメッセージを走査して、既知の望ましくないメッセージ１および２があるかどうかを確認している可能性がある。数分後に、ユーザＢのメッセージを走査して、望ましくないメッセージ１、２、３が存在するかどうかを確認することができる。３０分後にユーザＣは第４の望ましくないメッセージ４を発見する可能性があり、１時間後にユーザＡのメッセージをもう一度走査して、この場合は１、２、３、４が存在するかどうかを確認することができる。本発明は、アウトバウンド・メッセージならびにインバウンド・メッセージの走査に備えるものである。これは、でっちあげ、チェーン・レター、機密メッセージなど、１人のユーザから他の複数のユーザに転送されそうなタイプのメッセージについては特に有利である。転送される前に望ましくないアウトバウンド・メッセージを捕らえることは、多数の受信側になる可能性のあるものに対して送信された後でメッセージを処理することよりかなり効率のよいことである。

【００４７】本発明の第１の段階でメッセージから抽出され、本発明の第２の段階で重複メッセージまたは類似メッセージを認識するために後で使用されるシグニチャ・データを表す好ましいデータ構造については図３に示す。当業者であれば、本発明では多少精巧なデータ構造を使用できることが分かるだろう。望ましくないメッセージは実質的に類似したメッセージのセットにクラスタ化される。１つのクラスタ内には、原型と呼ばれる１つまたは複数の変形が存在する可能性がある。多くの場合、各クラスタは単一の原型のみを含むことになる。しかし、事情によっては（特に、複数の関連変形として現れる可能性があるでっちあげの場合）、１つのメッセージのわずかな変形を同じクラスタに属すものと見なすことは有用である可能性がある。１つのクラスタ内に複数の原型があることを見込んでおくと、同じシグニチャを使用して複数の異なる変形を検出することができる。この結果、記憶の効率が高まり、走査速度がある程度高くなり、また、新しい変形がそれとして認識される可能性も高くなる。さらに、本発明のシグニチャ抽出データの精巧さは、変形を検出することと偽陽性を低減することとの間でトレードオフが行われるようにシステムを調整する際のフレキシビリティに備えるものである。

【００４８】本発明の一実施の形態のシグニチャ・データベースは、原型Clusterのセットからなり、それぞれは固有のClusterID識別子によって区別される。各Cluster３００は２つの基本コンポーネントを有する。第１のコンポーネントはSigList３０２である。SigList３０２はSigData要素３０４のリストであり、そのそれぞれは原型Cluster３００のメンバー内で検出された特定の文字シーケンスに関する情報を含む。３つのSigData要素であるSigData1、SigData2、SigData3を示す。SigList３０２内の各SigData要素３０４は２つの部分を含む。たとえば、SigData2だけを展開する。SigData2 ３０４の第１の部分であるSig2 ３０６は、メッセージ・スキャナによってサーチされる相対的に短いテキスト・パターンである。第２の部分であるRegionList2 ３０８はSig2 ３０６に関連するRegionData要素３１０のリストであり、そのそれぞれはクラスタ内のすべての原型に含まれるより長い文字シーケンスに関する情報を含む。各RegionData要素３１０は３つの要素、すなわち、１）シグニチャの先頭から文字シーケンスの先頭までのバイト単位のオフセットであるBeginOffset３１２と、２）文字シーケンス内の文字数であるRegionLength３１４と、３）文字シーケンスのチェックサムであるＣＲＣ３１６である。

【００４９】各Cluster３００の第２のコンポーネントはArchetypeList３１８である。ArchetypeList３１８はArchetypeData要素３２０のリストであり、そのそれぞれは特定の原型に関するデータを含む。特に、各ArchetypeData要素３２０は、１）そのフル・テキストを必要に応じて検索できるように原型メッセージの記憶コピーを指すポインタであるArchetypePtr３２２と、２）原型の本文から計算され、他のメッセージに対する全体的な

類似性を測定するために使用するデータのブロックであるHashBlock３２４と、３）CaseData要素３２８のリストであり、そのそれぞれが原型のコピーが受信され、ユーザによって望ましくないものとして報告された特定のインスタンスに関するデータを含むCaseList３２６とを含むことができる。特に、各CaseData要素３２８は、１）コピーの送信側のＩＤであるSendID３３０と、２）コピーを報告した受信側のＩＤであるRecvID３３２と、３）コピーが本来受信された時刻であるRecvTime３３４とを含むことができる。

【００５０】シグニチャ抽出

本発明のシグニチャ抽出段階の好ましい実施の形態は、その間に特定の以前は未知であった望ましくない（または機密の）メッセージを検出するための方法が導出され、複数ユーザからなるネットワークに広められるものであるが、図４に関連して説明する。本発明は、１人または複数のメール・ユーザを含む環境で使用することができる。メール・ユーザの数が増すにつれて、本発明の利点が増大する。ステップ４００では、第１の（アラート）ユーザがメッセージＭ１を受信する。このユーザは受信したメッセージＭ１を読み、それが広く流布されそうであり、歓迎されないものであると広く考えられそうなものである（またはそれが機密のものである）という意味で「望ましくないもの」であると考えた場合、そのユーザは、たとえば、ユーザ・インタフェース内の特別なボタンをクリックすることにより、メッセージＭ１は望ましくない（または機密の）ものとしてフラグを付けるべきであることをシステムに対して示す。任意選択で、第１に総称検出方法を使用して、ユーザがそのメッセージを望ましくないものとして識別するのを支援することができる。いずれの場合でも、ステップ４０２でそのメッセージに「望ましくないもの」としてのフラグを付けるべきであることをユーザがシステムに対して示した場合、ステップ４０４でメッセージＭ１のコピーを自動シグニチャ抽出手順に送信するかまたは入力するかあるいはその両方を行う。任意選択で、ステップ４０３では、望ましくないものとしてのメッセージの識別をいくつかの方法で確認することができる。この確認は、許可人間ユーザによって行うこともできる。これは、限界数のユーザ全員がそのメッセージに望ましくないものとしてのラベルを付けた後でのみ示される可能性がある。最後に、これは、個別の自動プロセス（たとえば、スパムを検出するための総称技法を使用するもの）によって行うこともできる。メッセージが望ましくないものであるという確認が行われた場合、この方法はステップ４０４に継続するはずである。メール・システム・ユーザ自身が望ましくないかまたは機密のメッセージを識別できるようにすることにより、集中オペレーション・センタにいる専門家への依存状態が回避される。

【００５１】ステップ４０４では、メッセージＭ１を走査して、マスタ・シグニチャ・データベースＤ１に含まれるシグニチャが存在するかどうかを確認する。ステップ４０５でメッセージＭ１がマスタ・シグニチャ・データベースＤ１内のシグニチャの少なくとも１つを含むと判断された場合、ステップ４０６でそのメッセージは、そのSigコンポーネントの１つに一致シグニチャを含む各Clusterに関連する各原型と比較して、Ｄ１内のいずれかの原型との一致が存在するかどうかを判定する。比較の好ましい方法は、そのメッセージ関するHashBlockを計算することと、このHashBlockを各候補原型に関するHashBlockと比較することである。正確な原型一致が見つかった場合（たとえば、ハッシュブロック距離がゼロになると計算された場合）、一致候補のArchetypePtr３２２を使用して、そのフル・テキストを検索する。最後に、その原型およびメッセージのフル・テキストが、そのメッセージを原型のインスタンスと見なせるほど十分類似していると見なされた場合、ステップ４０８で関連CaseData情報３２８をメッセージから抽出し、その原型用のＤ１内のCaseList３２６に追加する。次に制御はステップ４１８に移行する。しかし、ステップ４０６で正確な原型一致が見つからないかまたはメッセージのフル・テキストがその原型のフル・テキストと十分に類似していると判定されなかった場合、ステップ４１０で新しい原型が既存の原型のクラスタと十分に類似しているかどうかについて判定を行い、十分類似している場合はどのクラスタであるかという判定を行う。好ましいことに、そのSigコンポーネントの１つに一致シグニチャを含む各Clusterごとに、BeginOffset３１２およびRegionLength３１４によって示される領域のチェックサムを計算することにより、そのSig３０６に関連するRegionList３０８内の各RegionData要素３１０をメッセージＭ１と比較し、そのメッセージ内のその領域のチェックサムがＣＲＣ３１６に記憶されている値に等しい場合に一致を宣言する。最長RegionLength３１４を有する一致領域を各Clusterごとに決定する。すべてのCluster内の最長RegionLength３１４が少なくとも指定の限界長に等しい場合、最長RegionLength３１４を有するClusterは、新しい原型を追加すべき原型クラスタとして識別される。したがって、ステップ４１２では、原型データを計算し、それを（すべての副構造が必須情報で充填された）新しいArchetypeData要素としてこのClusterのArchetypeListに追加する。

【００５２】任意選択で、ステップ４１４では、そのクラスタへの新しい原型の追加を反映するために、ClusterのSigList３０２を再計算することができる。突合せアルゴリズム（接尾部アレイ・ルーチンなど）を使用してすべての原型間で見つかった１つまたは複数の文字シーケンスを識別することができ、図５に関連して以下に詳述するSigListデータの導出は、メッセージ本文全体ではなく、一般的に発生する文字シーケンスのセットにの

み適用することができる。この方法はステップ４１８に継続する。

【００５３】ステップ４０５でメッセージＭ１がマスタ・シグニチャ・データベースＤ１内のシグニチャのいずれも含まないと判断された場合またはステップ４１０でいかなる原型クラスタも新しい原型に十分接近していないと判断された場合、この方法はステップ４１６に継続する。ステップ４１６でメッセージＭ１用に新しい原型クラスタを作成し、必須情報を含む単一のArchetypeData要素を作成してArchetypeList内に入れ、シグニチャと関連データのセットを計算してSigList内に入れる。最後に、原型Clusterにその固有のClusterIDを割り当て、そのClusterをマスタ・シグニチャ・データベースＤ１に追加する。SigList内のシグニチャは、他のメッセージで見つかりそうもない文字シーケンスを選択する自動シグニチャ抽出手順によって自動的に計算される。この手順の好ましい方法に関する詳細については図５に関連して以下に示す。シグニチャは、そのメッセージ自体またはそのメッセージの前処理バージョンで見つかる複数文字からなるシーケンスまたはより一般的には複数文字からなるパターンで構成することができる。これには、メッセージ全体またはその一部分あるいはその両方のチェックサム、そのメッセージの１回または複数回の変換から導出されるチェックサムまたはその他の圧縮データ・ストリングなどの追加情報が付随する可能性がある。この追加情報は、図３に示す各シグニチャに関連するRegionList３０８に記憶することができる。

【００５４】最後に、ステップ４１８では、ステップ４０８、４１４、または４１６でマスタ・シグニチャ・データベースＤ１に対して適用された更新を反映するために、１つまたは複数の個別ユーザ・ノードに対応するローカル・シグニチャ・データベースを更新する。これは、ローカル・データベースがマスタ・シグニチャ・データベースの正確な複製であることを保証するための標準的なデータベース更新または複写技法を使用するか、あるいは異なるローカル・シグニチャ・データベース間で様々になる可能性のある１組の基準に応じてシグニチャおよび関連補助データを選択的に送信するかまたは選択的に受信して取り入れることによって達成することができる。

【００５５】SigListデータの導出

次に、ステップ４１４および４１６で使用する、所与の原型メッセージに関するSigListデータを抽出または計算するための手順の好ましい実施の形態について図５に関連して説明する。第１に、ステップ５００では、メッセージのコーパス内で選択した限界長より小さいかまたはそれに等しいすべてのバイト・シーケンスのオカレンスの回数を数える。好ましい実施の形態では、限界長は３であり、すなわち、１バイト、２バイト、３バイトのすべてのシーケンス（それぞれ１グラム、２グラム、３グラムという）のオカレンスの回数を数える。ステップ５０１では、数えたオカレンスの回数をｎグラム度数データベースに圧縮形式で記憶する。ｎグラム度数データベースは、わずか数メガバイトのデータベースを必要とする。このデータベースはそのユーザが受信したアーカイブ対象メッセージからなるコーパスから個別に各ユーザごとに計算することができ、あるいは汎用データベースは複数のユーザから選別された総称メッセージからなる標準的なコーパスから計算することもできる。この汎用データベースは、ユーザ全体にわたって配布することもできる。また、このデータベースは定期的に更新することもできる。データベースが本来生成される場所およびそれが更新される頻度に関する詳細は、シグニチャ抽出手順の残りのステップに何の関係もない。

【００５６】ステップ５０２では、そこからシグニチャを抽出すべきメッセージＭ２の本文を分離する。ステップ５０４では、すべての非英数字を除去し、すべての上段シフト文字をその下段シフト・バージョンで置き換えることにより（図６を参照）、抽出した本文を「不変」形式に変換する。次に、ステップ５０６では、典型的なメッセージで見つからない可能性が高い文字からなる１つまたは複数のシーケンスを識別する。この１つまたは複数のシーケンスは１つまたは複数のシグニチャを構成する。ありそうもない文字シーケンスの識別は、参照により本明細書に組み込まれ、１９９５年９月１９日に発行された「Methods and Apparatus forEvaluating and Extracting Signatures of Computer Viruses and Other Undesirable Software Entities」という名称の米国特許第５４５２４４２号（４４２特許）に記載された方法によって実施することができる。この方法は、本来、コンピュータ・ウィルス・シグニチャの自動抽出に適用されたものである。メッセージから取られる複数の候補シグニチャを選択し、ｎグラム度数データベースからのそれぞれのｎグラム統計について、各候補シグニチャが無作為の普通のメール・メッセージに現れる可能性を推定するために４４２特許で見られる式を使用してそれらを結合する。普通のメール・メッセージに現れる可能性が最も低い１つまたは複数の候補シグニチャを選択する。

【００５７】ひとまとめにして考えると、ステップ５０２、５０４、５０６は、図３でSig３０６というラベルが付けられたテキスト・ストリング要素の導出を記述するものである。任意選択で、Sig３０６に関連するRegionData３１０のリストを計算することによって、偽陽性率をさらに低減することができる。これは、各導出シグニチャごとに以下の手順によってステップ５０８で達成することができる。それぞれがそのシグニチャを含む文字シーケンスからなる「領域」のシリーズを選択する。好ましい実施の形態では、このシリーズは、シグニチャ上のほぼ中心に位置し、シグニチャの長さの約２倍であ

る第１の領域と、第１の領域を含み、第１の領域のサイズのほぼ２倍である第２の領域と、シリーズ内の最終領域が変換済みのメッセージ本文全体からなるまでの以下同様の領域からなる。各領域ごとに、その領域の長さおよびその領域の文字シーケンスのチェックサムとともに、そのシグニチャの第１の文字からその第１の文字までのオフセット（通常は負の整数）を記録する。これらの３つの要素はその領域のRegionData３１０を構成する。チェックサムは、巡回冗長検査などの都合のよい方法を使用することができ、好ましいことに少なくとも３２ビットにしなければならない。

【００５８】HashBlockデータの導出

次に、ステップ４１２および４１６で要求されたように所与のメッセージに関するHashBlockデータを計算するための方法の好ましい実施の形態について説明する。第１に、メッセージ本文を変換する。この変換は、シグニチャ抽出の前にメッセージ本文に対して適用される変換（ステップ５０４）と同じにするかまたはそれとは異なるものにすることができる。たとえば、HashBlockを計算するために変換済みのメッセージ本文にブランク・スペースを保持すること以外は、これらの変換を同一のものにすることもできる。次に、変換済みのメッセージ本文を小さい個別ユニットに分割するが、これらのユニット同士はオーバラップする場合もあれば、オーバラップしない場合もある。たとえば、個別ユニットはいずれも連続５文字のシーケンス（オーバラップする）である場合もあれば、オーバラップしない「ワード」（ブランク・スペースによって区切られた個別ユニット）である場合もある。オーバラップしないユニットの方が好ましい。各個別ユニットごとに、ハッシュ関数はそのユニットを小さい整数のハッシュ値（たとえば、０～２５５の範囲内）にマッピングする。ハッシュ値カウントのアレイは保たれ、特定のハッシュ値が計算されるたびに、その値のカウントが１だけ増分される。カウント数の上限が１５に定められている場合あるいはそれがモジュロ１６で計算される（すなわち、１６で割ったときに、記録された数が実際の数の剰余になる）場合、各カウントごとに４ビットだけが必要であり、２５６通りのハッシュ値からなるアレイはちょうど１２８バイトのHashBlockとして表すことができる。ただし、変更の回数が多すぎなければ、このHashBlockはワードの追加、削除、再配置に対して相対的に鈍感になることに留意されたい。

【００５９】シグニチャ・データベースの剪定

マスタおよびローカルのシグニチャ・データベースが無制限に増大するのを防止するために、これらのデータベースを定期的に剪定して、最近のインスタンスがまったく報告されていないClusterデータを除去することができる。好ましいことに、周期的間隔で（たとえば、毎日）マスタ・シグニチャ・データベース内の各Clusterを検査する。そのクラスタ構造内のすべてのRecvTime要素３３４を現在時と比較し、いずれかの指定の日時より最近のものがまったくない場合、そのCluster全体をマスタ・シグニチャ・データベースから除去する。このクラスタの除去はすべてのローカル・シグニチャ・データベースに連絡され、このクラスタを含むものはどれでもそれを除去することができる。

【００６０】シグニチャ走査

本発明のシグニチャ走査段階では、１人または複数のユーザのメッセージを走査して、望ましくない（または機密の）ものとしてラベルが付けられた特定のメッセージが存在する可能性があるかどうかを確認する。数百人、数千人、または数百万人ものユーザを本発明によって保護することができるが、個別の「第２のユーザ」に焦点を合わせることが最も都合のよいことである。この走査手順ではローカル・シグニチャ・データベースを使用するが、これは新しい望ましくないメッセージが他のユーザによって発見されるにつれて引き続いて更新され、特定のユーザに固有のものにするかまたは複数ユーザによって共用することができる。この走査は、定期的に行うか、あるいはユーザによる要求または何らかの他の事象（最後の走査以降にローカル・シグニチャ・データベースが更新されたという通知など）に応答して行うことができる。さらに、この走査は、様々なユーザについて様々な時期に様々な状況で行うことができる。メッセージが電子メールである典型的な場合には、走査は好ましいことにユーザの受信箱内の項目にのみ適用されるが、ユーザがそのように希望する場合には他の指定のフォルダにも適用することができる。

【００６１】走査手順の好ましい実施の形態について図６に関連して説明する。ステップ６０２では、走査すべきメッセージＭ２の本文を抽出する。次にステップ６０４では、メッセージ本文をステップ５０４で適用されたものと同じ不変形式に変換する。ステップ６０６では、メッセージ本文の不変形式を走査して、ローカル・シグニチャ・データベースＤ２に含まれるシグニチャのいずれかとの正確な一致または接近した一致があるかどうかを確認するが、このデータベースＤ２は１つまたは複数のマスタ・シグニチャ・データベース内のClusterデータ構造の全部または一部から構築されたものである。いかなるシグニチャも見つからない場合、メッセージは望ましくない（または機密の）ものと見なされ、プロセスは終了する。

【００６２】しかし、ステップ６０６で１つまたは複数のシグニチャが見つかった場合、ステップ６０８で関連のRegionData要素３１０に含まれる補助情報を使用して、１つまたは複数の既知の望ましくないメッセージに対する一致度を査定する。具体的には、メッセージに現れる各シグニチャSig３０６ごとに、Sig３０６が現れるすべてのClusterを順に検討する。このような各Cluster３００ごとに、Sig３０６に関連するRegionList３０８

を検討する。第１に、走査したメッセージ内の対応する領域のチェックサムを計算することにより、最大RegionLength３１４を有するRegionData要素３１０を検査する。チェックサムがこのRegionData要素３１０に関するＣＲＣ３１６と一致する場合、RegionData要素３１０および関連のClusterIDをBestRegionDataElementsというリストに追加し、次のClusterを検討する。チェックサムが一致しない場合、次に長いRegionLength３１４を有するRegionData要素３１０を同じ方法で比較し、一致チェックサムが見つかるまで以下同様に比較する。RegionData要素３１０間に一致チェックサムがまったく存在しない場合、そのシグニチャ自体および関連のClusterIDをBestRegionDataElementsリストに追加し、次のClusterを検討する。

【００６３】ステップ６１０では、局所性保存ハッシュ関数を使用して、走査したメッセージのHashBlockを計算する。走査したメッセージのHashBlockを、ステップ６０６で見つかった一致シグニチャの１つを含む各ClusterのHashBlockと比較し、このような各Clusterごとに類似性を計算する。この類似性の計算では妥当な測定基準を使用することができる。２つのHashBlock（Ｈ１およびＨ２）に関する好ましい類似性測定基準では、それぞれを２５６要素のアレイとして扱い、それぞれの要素は４ビットとして表され、アレイ要素間の差の絶対値を合計する。すなわち、類似性Ｓは以下の式によって示される。

【数１】

$$S = \sum_{j=0}^{255} \left| H_{1j} - H_{2j} \right| \qquad (1)$$

【００６４】これは、アレイ要素の上限が１６に定められている場合である。その代わりとして、以下の式によって示される。

【数２】

$$S = \sum_{j=0}^{255} \left(\left(H_{1j} - H_{2j} + 16 \right) \bmod 16 \right) \qquad (2)$$

【００６５】これは、アレイ要素がモジュロ１６で記憶される場合である。

【００６６】ClusterIDおよび類似性ＳをHashBlockSimilarityというリストに追加し、ステップ６０６で見つかった一致シグニチャの１つを含むClusterがこれ以上存在しなくなるまで、次のClusterを検討する。

【００６７】ステップ６１２では、ステップ６０８から導出したBestRegionDataElementsリストと、ステップ６１０から導出したHashBlock類似性リストと、１組のユーザ・プリファレンスとを組み合わせて、一致の程度またはレベルを決定する。ユーザ・プリファレンスは、HashBlock類似性に関する１つまたは複数のしきい値と、RegionLength３１４に関する１つまたは複数のしきい値と、BestRegionDataElementsリストおよびHashBlockSimilarityリストで参照されるClusterのMsgDataコンポーネントの様々な態様に関する条件とで構成することができる。典型的な応用例では、ユーザ・プリファレンスは、選択した場合は高度なユーザによって指定変更が可能な何らかのデフォルト設定に設定することができる。

【００６８】明示的な一例として、完全、高、中、低という４通りの別々の一致レベルがあると想定する。この場合、ユーザ・プリファレンスの妥当なセットとしては以下のものが考えられる。完全と見なすべき一致レベルの場合、それに関するHashBlock類似性距離がゼロであり、そのCluster用のMsgList内の少なくとも２人のユーザがそのユーザと同じＥメール・ドメイン内のRecvID３３２を有するようなClusterが存在しなければならない。そうではない場合、高と見なすべき一致レベルの場合、それに関するHashBlock類似性距離が５未満であるかまたはBestRegionDataElements内の最長領域長が少なくとも５００文字であり、そのCluster用のMsgList内の少なくとも２人のユーザがそのユーザと同じＥメール・ドメイン内のRecvID３３２を有するようなClusterが存在しなければならない。そうではない場合、中と見なすべき一致レベルの場合、それに関する最長領域長が少なくとも１００文字であり、そのMsgList内に少なくとも２人の別々のユーザが存在し、ドメインまたは他の特徴に関する制限がまったくないようなClusterが存在しなければならない。そうではない場合、一致レベルは低と見なすべきである。

【００６９】ステップ６１４では、ステップ６１２で決定した一致レベルに対してそのユーザのプリファレンスのセット内のもう１組の規則を適用して、適切な応答を決定し、実施する。適切な応答としては、メッセージを自動的に削除すること、ユーザの受信箱内でのその外観を変更すること（たとえば、それに注釈を付けるかまたはカラー化することによる）、それを特別なフォルダ内に記憶することなどを含むことができる。たとえば、一致レベルが完全である場合、ユーザは、そのメールが自動的に削除されるはずであると示すことができ、一致レベルが高である場合、そのメールは特別な「推定スパム」フォルダ内に置かれるはずであり、一致レベルが中である場合、受信箱内に現れるメールの要約は緑色でカラー表示されるはずであり、メッセージ本文の前には、そのメッセージが望ましくないメールの既知のインスタンスに密接に関連すると思われる理由を示す簡単な説明が付けられるはずである。また、ユーザのプリファレンスは、その一致レベルにかかわらず、望ましくないものと見なすべきではない特定のメッセージ（その管理者またはその企業の最高経営責任者から送信されたものなど）も指定することができる。

【００７０】任意選択で、望ましくないメッセージが発見された場合、ステップ６１６ではその望ましくないメ

ッセージの新しいインスタンスに関する情報によってマスタ・シグニチャ・データベースを更新することができる。この更新は、発見時に行われる場合もあれば、あるいはそのメッセージが望ましくないものであることをユーザが確認した後でのみ行われる場合もある。たとえば、完全な一致の場合、この情報は、望ましくないメッセージに関するCaseData３２８（すなわち、送信側および受信側のＩＤならびに受信時刻）で構成することができる。この情報は、ローカルで抽出して、マスタ・シグニチャ・データベースの所在地に送信することもでき、そこに取り入れられるはずである。一致レベルが高または中の場合、メッセージ全体がマスタ・シグニチャ・データベースの所在地に送信される可能性があり、それがステップ４０４でシグニチャ抽出段階に入り、そこで新しい原型を作成してそれを適切な原型クラスタに入れようという試みが行われるはずである。

【００７１】まとめとして、本発明の構成に関して以下の事項を開示する。

【００７２】（１）複数ユーザからなるネットワーク内で電子メッセージの望ましくない送信または受信を妨害する方法において、少なくとも１つの特定の電子メッセージの送信または受信が望ましくないものであることを判定するステップと、前記少なくとも１つの特定の電子メッセージまたはその変形の検出を可能にする検出データを自動的に抽出するステップと、少なくとも１人のユーザからの１つまたは複数のインバウンド・メッセージまたはアウトバウンド・メッセージあるいはその両方を走査して前記少なくとも１つの特定の電子メッセージまたはその変形が存在するかどうかを確認するステップと、前記走査ステップに応答して適切なアクションを実行するステップとを含む方法。

（２）抽出した前記検出データを記憶するステップをさらに含む、上記（１）に記載の方法。

（３）前記判定ステップが、前記少なくとも１つの特定の電子メッセージの増殖が望ましくないという通知を受信するステップを含む、上記（１）に記載の方法。

（４）前記受信ステップが、前記少なくとも１つの特定の電子メッセージを望ましくないものまたは機密のものとして識別する信号をアラート・ユーザから受信するステップを含む、上記（３）に記載の方法。

（５）前記少なくとも１つの特定の電子メッセージが前記アラート・ユーザの受信箱に受信される、上記（４）に記載の方法。

（６）前記受信ステップが、前記特定の電子メッセージに望ましくないものとしてのフラグを付けるべきであることを示すための識別子を前記アラート・ユーザに提供するステップを含む、上記（４）に記載の方法。

（７）前記提供ステップが、電子メッセージが望ましくないものであるという識別を援助するために総称検出器を提供するステップを含む、上記（６）に記載の方法。

（８）前記抽出ステップが、前記少なくとも１つの特定の電子メッセージからシグニチャ情報を抽出するステップを含む、上記（２）に記載の方法。

（９）前記記憶ステップが、前記走査ステップに応答して、前記少なくとも１つの特定の電子メッセージに関する情報を前記シグニチャ情報に追加するステップを含む、上記（８）に記載の方法。

（１０）前記抽出ステップが、前記少なくとも１つの特定の電子メッセージからシグニチャを抽出するステップを含む、上記（２）に記載の方法。

（１１）前記記憶ステップが、前記シグニチャを少なくとも１つのシグニチャ・データベースに記憶するステップを含む、上記（１０）に記載の方法。

（１２）前記シグニチャ・データベースが複数のシグニチャ・クラスタを含み、各クラスタが実質的に類似した電子メッセージに対応するデータを含む、上記（１１）に記載の方法。

（１３）前記シグニチャ・クラスタのそれぞれが、走査情報を有する文字シーケンス・コンポーネントと、特定のシグニチャ変形に関する識別情報を有する原型コンポーネントとを含む、上記（１２）に記載の方法。

（１４）前記走査情報が、特定の電子メッセージに関するサーチ文字シーケンスと、前記クラスタ内に表されるすべての電子メッセージに関する拡張文字シーケンス情報とを含み、前記識別情報が、特定のシグニチャ変形に関連する電子メッセージのフル・テキスト記憶コピーを指すポインタと、前記電子メッセージのハッシュブロックと、前記電子メッセージのコピーが受信され、その増殖が望ましくないものとしてアラート・ユーザによって報告された特定のインスタンスに対応するアラート・データとを含む、上記（１３）に記載の方法。

（１５）前記抽出ステップおよび前記走査ステップが、複数ユーザからなる前記ネットワークの全域で同時かつ非同期に行われる、上記（２）に記載の方法。

（１６）前記走査ステップの前に、前記少なくとも１つの特定の電子メッセージが望ましくないものであることを確認するステップをさらに含む、上記（４）に記載の方法。

（１７）前記確認ステップが、総称検出技法により前記少なくとも１つの特定の電子メッセージが望ましくないものであることを確認するステップを含む、上記（１６）に記載の方法。

（１８）前記確認ステップが、前記少なくとも１つの特定の電子メッセージが望ましくないものであることを所定の限界数のユーザが通知することを必要とするステップを含む、上記（１６）に記載の方法。

（１９）前記抽出ステップが、前記特定の電子メッセージを走査して前記少なくとも１つのシグニチャ・データベース内のシグニチャがあるかどうかを確認するステップと、前記走査ステップで一致シグニチャを検出したこ

29

とに応答して、前記一致シグニチャを一致クラスタ内の各メッセージ変形と比較するステップとを含む、上記（１４）に記載の方法。

（２０）前記比較ステップが、前記特定の電子メッセージに関するハッシュブロックを計算するステップと、計算した前記ハッシュブロックを各原型コンポーネントの前記識別情報内の変形ハッシュブロックと比較するステップとを含む、上記（１９）に記載の方法。

（２１）正確な変形ハッシュブロック一致が検出された場合に、前記ポインタを使用して前記変形一致のフル・テキスト記憶コピーを検索するステップと、前記変形一致のフル・テキスト記憶コピーと前記特定の電子メッセージのフル・テキストが前記特定の電子メッセージを前記変形のインスタンスと見なすのに十分なほど類似していると見なされた場合に、前記特定の電子メッセージからアラート・データを抽出し、それを前記変形一致に関するアラート・データに追加するステップと、正確な変形ハッシュブロック一致が検出されないかまたは前記特定の電子メッセージのフル・テキストが前記データベース内の前記変形のいずれかと十分に類似していないと判断された場合に、前記特定の電子メッセージがいずれかの既存のクラスタと十分に類似しているかどうかを判定するステップと、前記特定の電子メッセージがある既存のクラスタと十分に類似している場合に、特定の電子メッセージに関連する新しい識別情報を計算するステップと、前記特定の電子メッセージがある既存のクラスタと十分に類似していると判定されない場合に、前記特定の電子メッセージに関する新しいクラスタを作成するステップとをさらに含む、上記（２０）に記載の方法。

（２２）前記判定ステップが、各クラスタの前記拡張文字シーケンス情報に示された前記特定の電子メッセージの領域のチェックサムを計算するステップと、計算した前記チェックサムを各クラスタの前記拡張文字シーケンス情報内の記憶チェックサムと比較するステップとを含む、上記（２１）に記載の方法。

（２３）シグニチャ一致がまったく検出されない場合に、前記特定の電子メッセージに関する新しいクラスタを作成するステップをさらに含む、上記（１９）に記載の方法。

（２４）前記拡張文字シーケンス情報が、開始オフセット・フィールドと、領域長フィールドと、ＣＲＣフィールドとを含み、前記方法が、各クラスタごとに、最長領域長を有する一致領域を決定するステップと、すべてのクラスタのうちで最長領域長が少なくとも指定の限界長と等しい場合に、最長領域長クラスタを特定の電子メッセージ原型が追加される原型クラスタとして識別するステップとをさらに含む、上記（２２）に記載の方法。

（２５）識別した前記クラスタの走査情報を再計算するステップをさらに含む、上記（２３）に記載の方法。

（２６）前記アラート・データが、コピーが本来受信された時刻を有する受信時刻フィールドを含み、前記方法が、各シグニチャ・クラスタのすべての変形の受信時刻フィールドを現在時と定期的に比較するステップと、前記受信時刻フィールドのいずれも所定の日時より最近のものではないシグニチャ・クラスタを除去するステップとをさらに含む、上記（１４）に記載の方法。

30

（２７）前記走査ステップが、メッセージ本文を抽出するステップと、前記メッセージ本文を不変形式に変換するステップと、前記不変形式を走査して前記検出データに対する正確な一致または接近した一致があるかどうかを確認するステップと、各一致ごとに一致のレベルを決定するステップとを含む、上記（１）に記載の方法。

（２８）前記実行ステップが、前記少なくとも１つの特定の電子メッセージまたはその変形の存在を発見したときに適切なアクションを実行するステップを含む、上記（１）に記載の方法。

（２９）前記実行ステップが、前記少なくとも１つの特定の電子メッセージまたはその変形に望ましくないものまたは機密のものとしてのラベルを付けるステップを含む、上記（２８）に記載の方法。

（３０）前記実行ステップが、前記少なくとも１つの特定の電子メッセージまたはその変形を除去するステップを含む、上記（２８）に記載の方法。

（３１）前記実行ステップが、１つまたは複数のユーザ・プリファレンスに応答して、決定した一致のレベルごとに適切なアクションを実行するステップを含む、上記（２７）に記載の方法。

（３２）前記決定ステップが、各一致ごとに最長領域一致を検出するステップと、走査したメッセージのハッシュブロックと抽出した検出データのそれぞれのハッシュブロックの間のハッシュブロックの類似性を計算するステップと、１つまたは複数のユーザ・プリファレンスを受信するステップと、前記検出ステップ、前記計算ステップ、および前記受信ステップに応答して、一致のレベルを決定するステップとを含む、上記（２７）に記載の方法。

（３３）複数ユーザからなるネットワーク内で電子メッセージの望ましくない送信または受信を妨害するための方法ステップを実行するためにマシンによって実行可能な命令からなるプログラムを具体的に実施する、マシンによって読取り可能なプログラム記憶装置において、前記方法が、少なくとも１つの特定の電子メッセージの送信または受信が望ましくないものであることを判定するステップと、前記少なくとも１つの特定の電子メッセージまたはその変形の検出を可能にする検出データを自動的に抽出するステップと、少なくとも１人のユーザからの１つまたは複数のインバウンド・メッセージまたはアウトバウンド・メッセージあるいはその両方を走査して前記少なくとも１つの特定の電子メッセージまたはその変形が存在するかどうかを確認するステップと、前記走

査ステップに応答して適切なアクションを実行するステップとを含むプログラム記憶装置。

（３４）複数ユーザからなるネットワーク内で電子メッセージの望ましくない送信または受信を妨害するためのシステムにおいて、少なくとも１つの特定の電子メッセージの送信または受信が望ましくないものであることを判定する手段と、前記少なくとも１つの特定の電子メッセージまたはその変形の検出を可能にする検出データを自動的に抽出する手段と、少なくとも１人のユーザからの１つまたは複数のインバウンド・メッセージまたはアウトバウンド・メッセージあるいはその両方を走査して前記少なくとも１つの特定の電子メッセージまたはその変形が存在するかどうかを確認する手段と、前記走査手段に応答して適切なアクションを実行する手段とを含むシステム。

（３５）抽出した前記検出データを記憶する手段をさらに含む、上記（３４）に記載のシステム。

（３６）前記判定手段が、前記少なくとも１つの特定の電子メッセージの増殖が望ましくないという通知を受信する手段を含む、上記（３４）に記載のシステム。

（３７）前記受信手段が、前記少なくとも１つの特定の電子メッセージを望ましくないものまたは機密のものとして識別する信号をアラート・ユーザから受信する手段を含む、上記（３６）に記載のシステム。

（３８）前記少なくとも１つの特定の電子メッセージが前記アラート・ユーザの受信箱に受信される、上記（３７）に記載のシステム。

（３９）前記受信手段が、前記特定の電子メッセージに望ましくないものとしてのフラグを付けるべきであることを示すための識別子を前記アラート・ユーザに提供する手段を含む、上記（３７）に記載のシステム。

（４０）前記提供手段が、電子メッセージが望ましくないものであるという識別を援助するために総称検出器を提供する手段を含む、上記（３９）に記載のシステム。

（４１）前記抽出手段が、前記少なくとも１つの特定の電子メッセージからシグニチャ情報を抽出する手段を含む、上記（３５）に記載のシステム。

（４２）前記記憶手段が、前記走査手段に応答して、前記少なくとも１つの特定の電子メッセージに関する情報を前記シグニチャ情報に追加する手段を含む、上記（４１）に記載のシステム。

（４３）前記抽出手段が、前記少なくとも１つの特定の電子メッセージからシグニチャを抽出する手段を含む、上記（３５）に記載のシステム。

（４４）前記記憶手段が、前記シグニチャを少なくとも１つのシグニチャ・データベースに記憶する手段を含む、上記（４３）に記載のシステム。

（４５）前記シグニチャ・データベースが複数のシグニチャ・クラスタを含み、各クラスタが実質的に類似した電子メッセージに対応するデータを含む、上記（４４）に記載のシステム。

（４６）前記シグニチャ・クラスタのそれぞれが、走査情報を有する文字シーケンス・コンポーネントと、特定のシグニチャ変形に関する識別情報を有する原型コンポーネントとを含む、上記（４５）に記載のシステム。

（４７）前記走査情報が、特定の電子メッセージに関するサーチ文字シーケンスと、前記クラスタ内に表されるすべての電子メッセージに関する拡張文字シーケンス情報とを含み、前記識別情報が、特定のシグニチャ変形に関連する電子メッセージのフル・テキスト記憶コピーを指すポインタと、前記電子メッセージのハッシュブロックと、前記電子メッセージのコピーが受信され、その増殖が望ましくないものとしてアラート・ユーザによって報告された特定のインスタンスに対応するアラート・データとを含む、上記（４６）に記載のシステム。

（４８）前記抽出手段および前記走査手段が、複数ユーザからなる前記ネットワークの全域で同時かつ非同期に処理する、上記（３５）に記載のシステム。

（４９）前記少なくとも１つの特定の電子メッセージが望ましくないものであることを確認する手段をさらに含む、上記（３５）に記載のシステム。

（５０）前記確認手段が、総称検出技法により前記少なくとも１つの特定の電子メッセージが望ましくないものであることを確認する手段を含む、上記（４９）に記載のシステム。

（５１）前記確認手段が、前記少なくとも１つの特定の電子メッセージが望ましくないものであることを所定の限界数のユーザが通知することを必要とする手段を含む、上記（４９）に記載のシステム。

（５２）前記抽出手段が、前記特定の電子メッセージを走査して前記少なくとも１つのシグニチャ・データベース内のシグニチャがあるかどうかを確認する手段と、前記走査手段で一致シグニチャを検出したことに応答して、前記一致シグニチャを一致クラスタ内の各メッセージ変形と比較する手段とを含む、上記（４７）に記載のシステム。

（５３）前記比較手段が、前記特定の電子メッセージに関するハッシュブロックを計算する手段と、計算した前記ハッシュブロックを各原型コンポーネントの識別情報内の変形ハッシュブロックと比較する手段とを含む、上記（５２）に記載のシステム。

（５４）正確な変形ハッシュブロック一致が検出された場合に、前記ポインタを使用して前記変形一致のフル・テキスト記憶コピーを検索する手段と、前記変形一致のフル・テキスト記憶コピーと前記特定の電子メッセージのフル・テキストが前記特定の電子メッセージを前記変形のインスタンスと見なすのに十分なほど類似していると見なされた場合に、前記特定の電子メッセージからアラート・データを抽出し、それを前記変形一致に関するアラート・データに追加する手段と、正確な変形ハッシ

ュブロック一致が検出されないかまたは前記特定の電子メッセージのフル・テキストが前記データベース内の前記変形のいずれかと十分に類似していないと判断された場合に、前記特定の電子メッセージがいずれかの既存のクラスタと十分に類似しているかどうかを判定する手段と、前記特定の電子メッセージがある既存のクラスタと十分に類似している場合に、特定の電子メッセージに関連する新しい識別情報を計算する手段と、前記特定の電子メッセージが既存のクラスタと十分に類似していると判定されない場合に、前記特定の電子メッセージに関する新しいクラスタを作成する手段とをさらに含む、上記（５３）に記載のシステム。

（５５）前記判定手段が、各クラスタの前記拡張文字シーケンス情報に示された前記特定の電子メッセージの領域のチェックサムを計算する手段と、計算した前記チェックサムを各クラスタの前記拡張文字シーケンス情報内の記憶チェックサムと比較する手段とを含む、上記（５４）に記載のシステム。

（５６）シグニチャ一致がまったく検出されない場合に、前記特定の電子メッセージに関する新しいクラスタを作成する手段をさらに含む、上記（５２）に記載のシステム。

（５７）前記拡張文字シーケンス情報が、開始オフセット・フィールドと、領域長フィールドと、ＣＲＣフィールドとを含み、前記システムが、各クラスタごとに、最長領域長を有する一致領域を決定する手段と、すべてのクラスタのうちで最長領域長が少なくとも指定の限界長と等しい場合に、最長領域長クラスタを特定の電子メッセージ原型が追加される原型クラスタとして識別する手段とをさらに含む、上記（５５）に記載のシステム。

（５８）識別した前記クラスタの走査情報を再計算する手段をさらに含む、上記（５６）に記載のシステム。

（５９）前記アラート・データが、コピーが本来受信された時刻を有する受信時刻フィールドを含み、前記システムが、各シグニチャ・クラスタのすべての変形の受信時刻フィールドを現在時と定期的に比較する手段と、前記受信時刻フィールドのいずれも所定の日時より最近のものではないシグニチャ・クラスタを除去する手段とをさらに含む、上記（４７）に記載のシステム。

（６０）前記走査手段が、メッセージ本文を抽出する手段と、前記メッセージ本文を不変形式に変換する手段と、前記不変形式を走査して前記検出データに対する正確な一致または接近した一致があるかどうかを確認する手段と、各一致ごとに一致のレベルを決定する手段とを含む、上記（３４）に記載のシステム。

（６１）前記実行手段が、前記少なくとも１つの特定の電子メッセージまたはその変形の存在を発見したときに適切なアクションを実行する手段を含む、上記（３４）に記載のシステム。

（６２）前記実行手段が、前記少なくとも１つの特定の電子メッセージまたはその変形に望ましくないものまたは機密のものとしてのラベルを付ける手段を含む、上記（６１）に記載のシステム。

（６３）前記実行手段が、前記少なくとも１つの特定の電子メッセージまたはその変形を除去する手段を含む、上記（６１）に記載のシステム。

（６４）前記実行手段が、１つまたは複数のユーザ・プリファレンスに応答して、決定した一致のレベルごとに適切なアクションを実行する手段を含む、上記（６０）に記載のシステム。

（６５）前記決定手段が、各一致ごとに最長領域一致を検出する手段と、走査したメッセージのハッシュブロックと抽出した検出データのそれぞれのハッシュブロックの間のハッシュブロックの類似性を計算する手段と、１つまたは複数のユーザ・プリファレンスを受信する手段と、前記検出手段、前記計算手段、および前記受信手段に応答して、一致のレベルを決定する手段とを含む、上記（６０）に記載のシステム。

【図面の簡単な説明】

【図１】本発明の教示を実施するためのコンピュータ・システムのブロック図である。

【図２】本発明の一実施の形態が適用されるシステム環境の概略図である。

【図３】本発明の一実施の形態のシグニチャ・データ構造の概略図である。

【図４】本発明の一実施の形態のシグニチャ抽出段階の流れ図である。

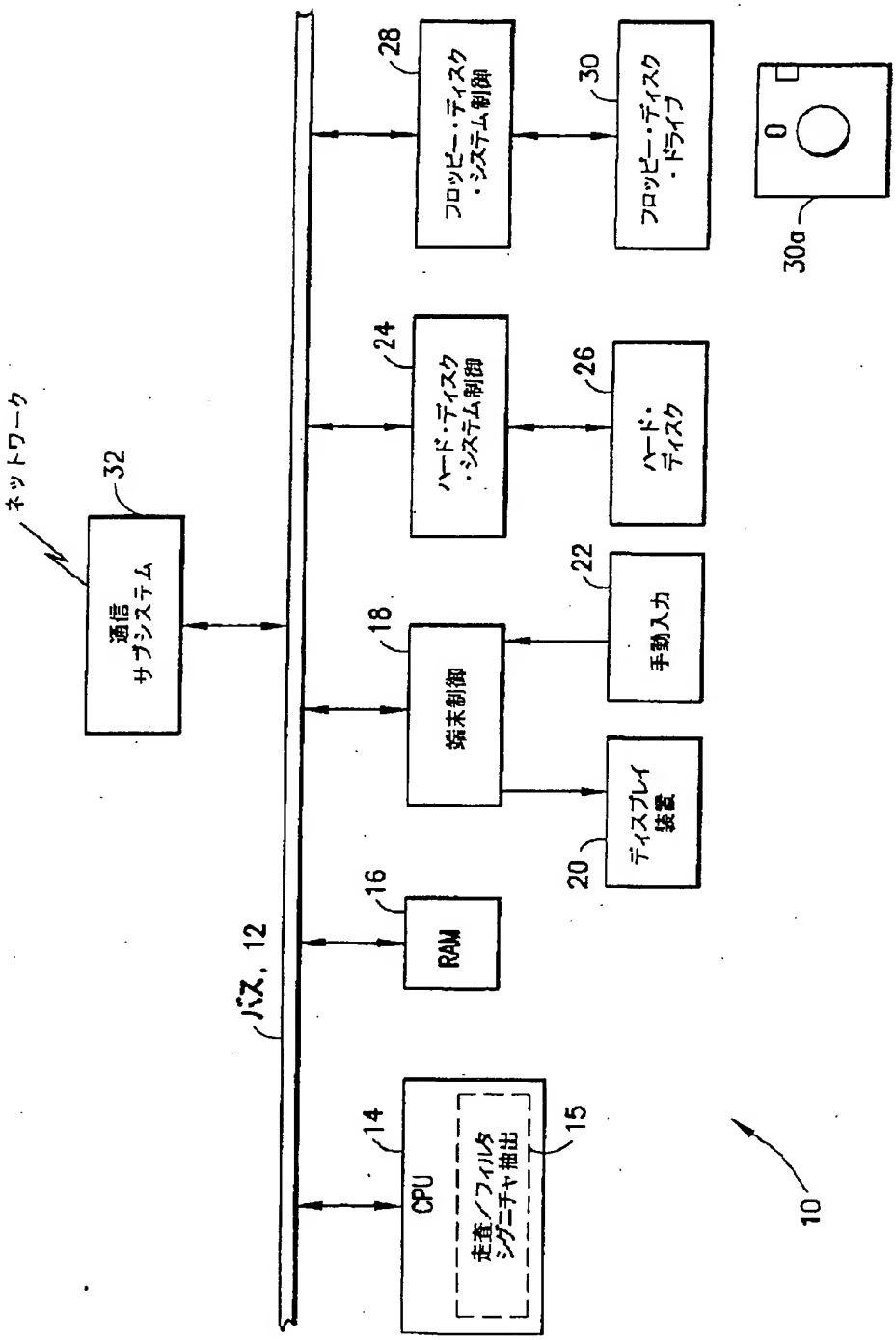
【図５】本発明の一実施の形態のシグニチャ抽出手順の詳細の流れ図である。

【図６】本発明の一実施の形態のシグニチャ走査段階の流れ図である。

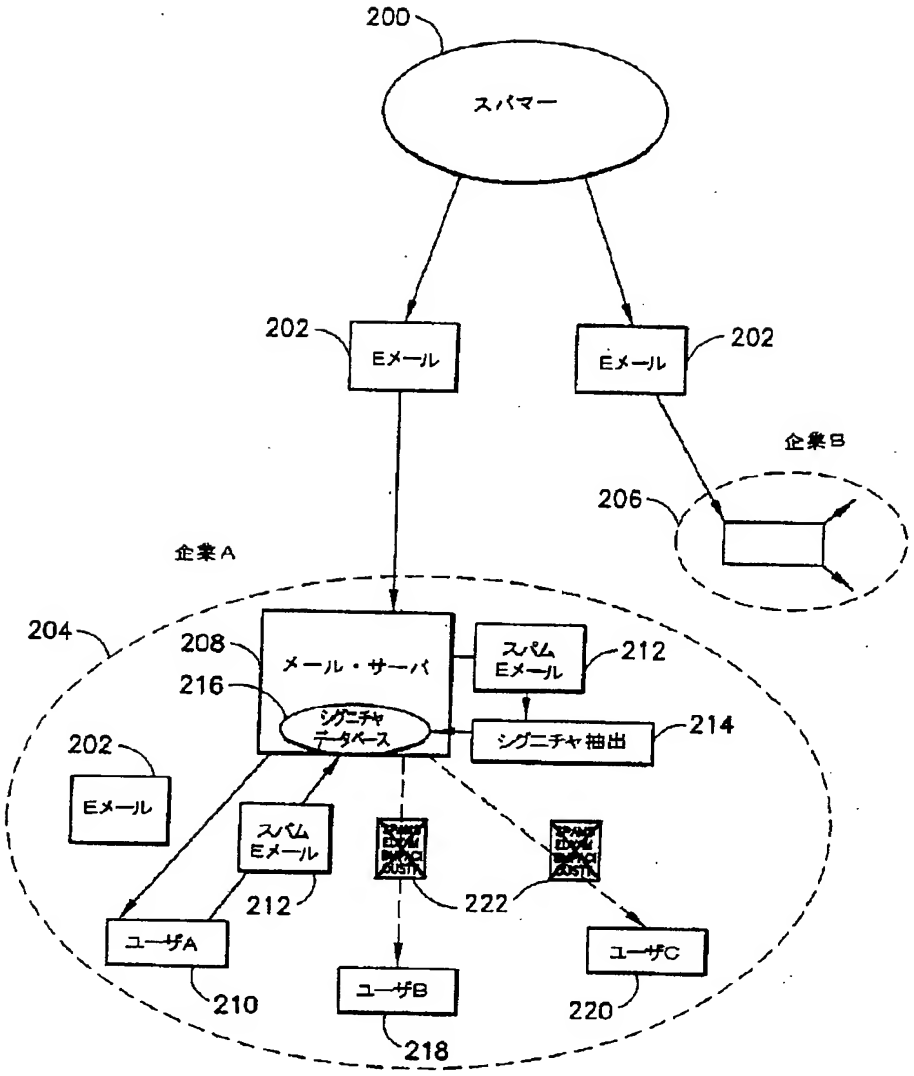
【符号の説明】

１０　システム
１２　システム・バス
１４　中央演算処理装置（ＣＰＵ）
１５　シグニチャ抽出モジュールおよび走査／フィルタ・モジュール
１６　ランダム・アクセス・メモリ（ＲＡＭ）
１８　端末制御サブシステム
２０　ディスプレイ装置
２２　手動入力装置
２４　ハード・ディスク制御サブシステム
２６　回転固定ディスクまたはハード・ディスク
２８　フロッピー・ディスク制御サブシステム
３０　フロッピー・ディスク・ドライブ
３２　通信サブシステム

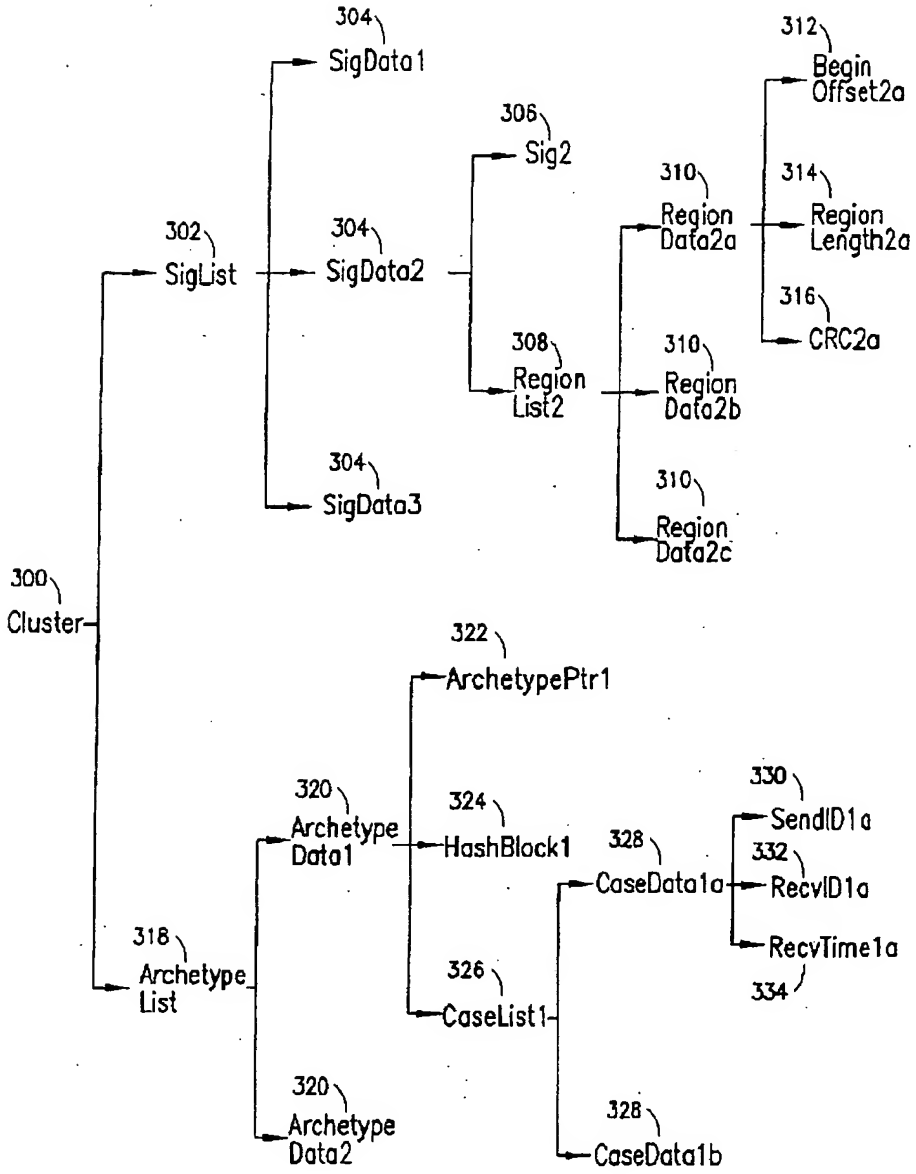
【図１】


ネットワーク
32
通信
サブシステム
バス, 12
CPU
走査／フィルタ
シグニチャ抽出
14
15
RAM
16
端末制御
18
ディスプレイ
装置
20
手動入力
22
ハード・ディスク
・システム制御
24
ハード・
ディスク
26
フロッピー・ディスク
・システム制御
28
フロッピー・ディスク
・ドライブ
30
30a
10

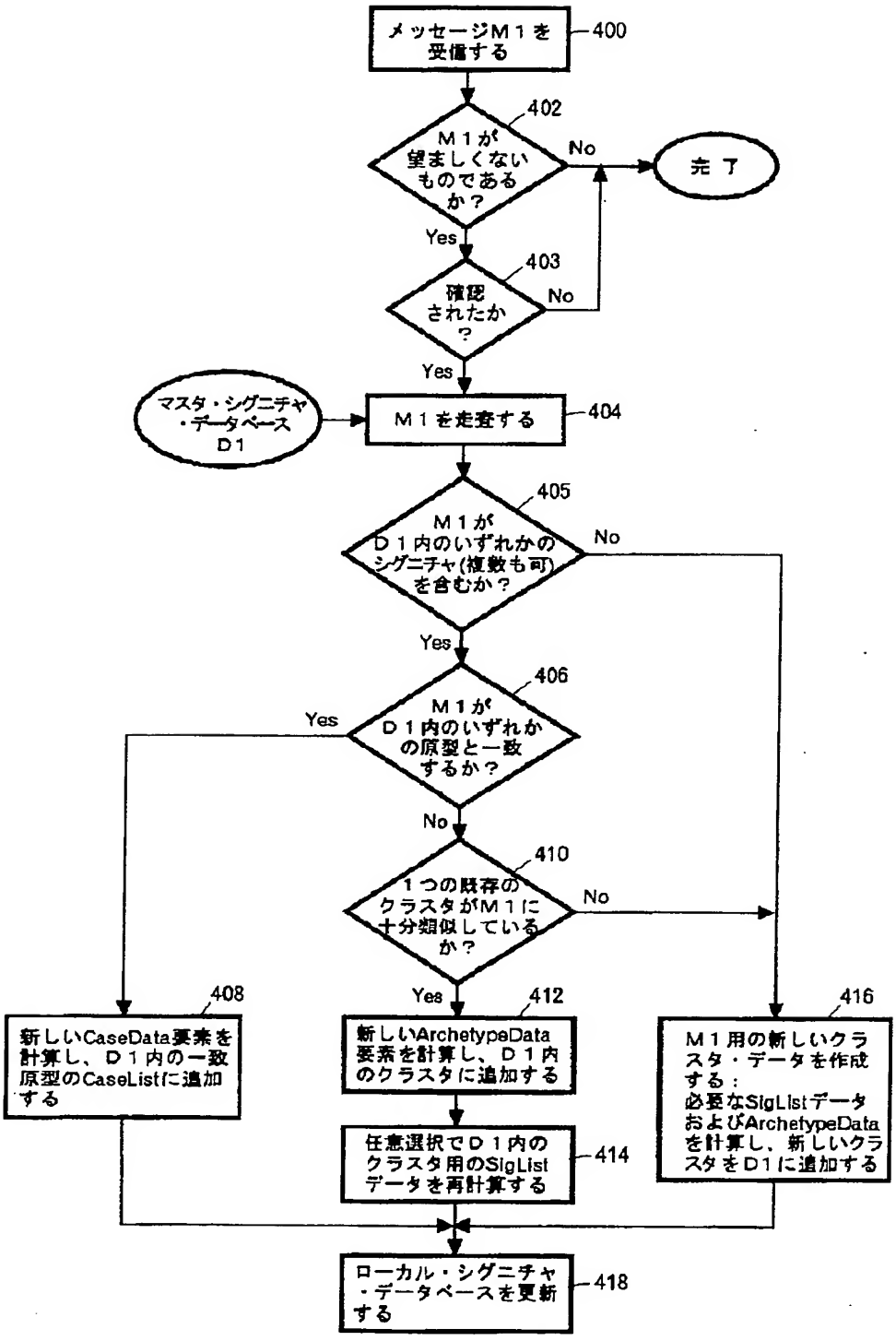
【図２】


200
スパマー
202
Eメール
Eメール
202
企業B
206
企業A
204
208
メール・サーバ
スパム
Eメール
212
216
シグニチャ
データベース
シグニチャ抽出
214
202
Eメール
スパム
Eメール
212
222
ユーザA
210
ユーザB
218
ユーザC
220

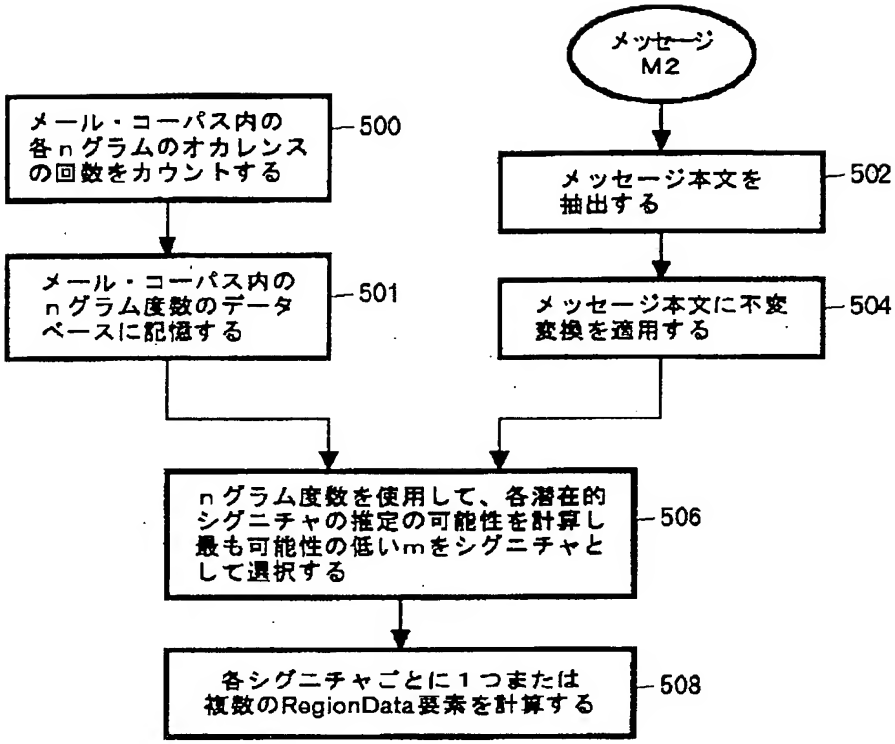
【図３】


300
Cluster
302
SigList
304
SigData1
304
SigData2
304
SigData3
306
Sig2
308
Region List2
310
Region Data2a
310
Region Data2b
310
Region Data2c
312
Begin Offset2a
314
Region Length2a
316
CRC2a
318
Archetype List
320
Archetype Data1
320
Archetype Data2
322
ArchetypePtr1
324
HashBlock1
326
CaseList1
328
CaseData1a
328
CaseData1b
330
SendID1a
332
RecvID1a
334
RecvTime1a

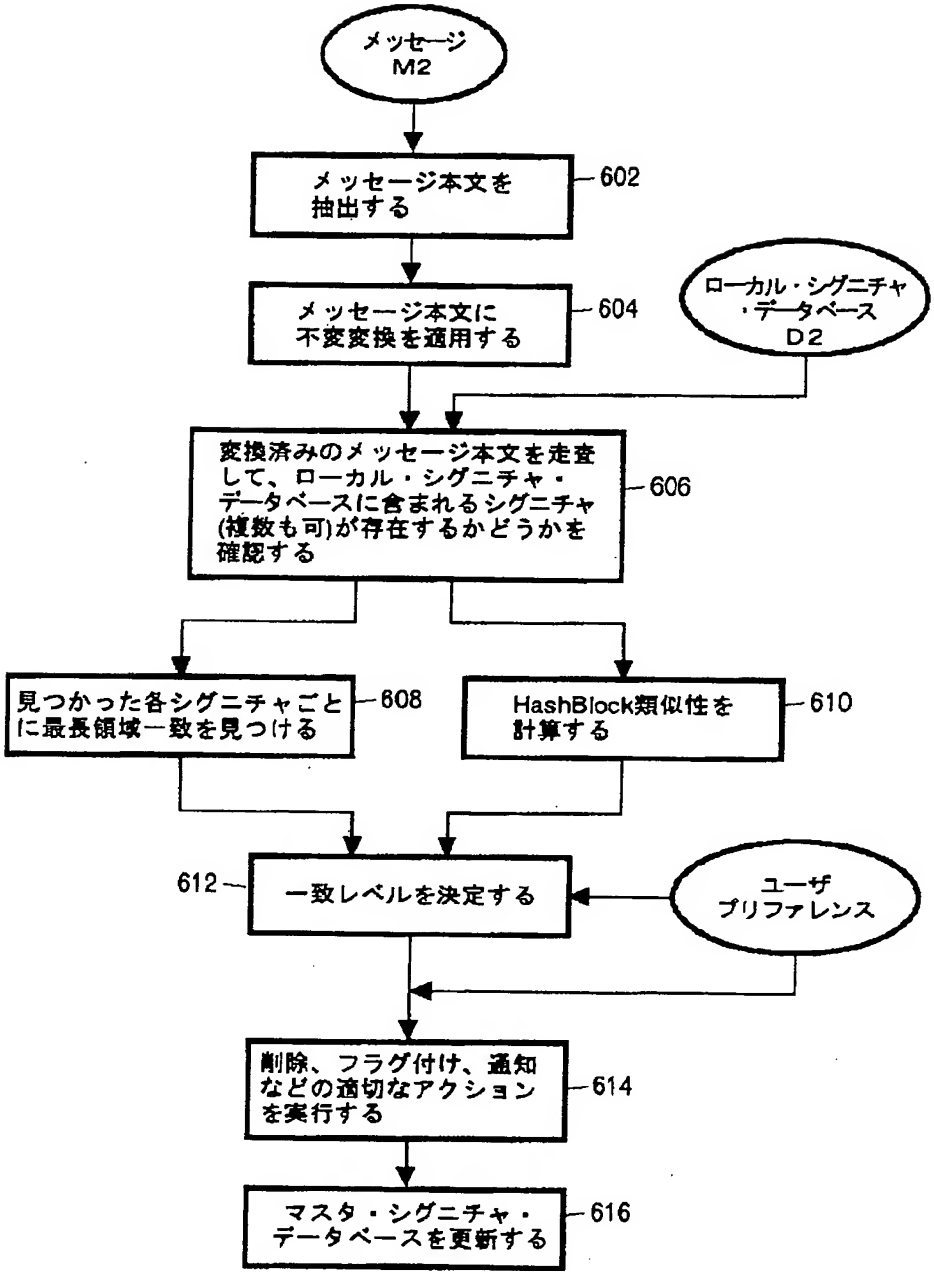
【図４】


メッセージＭ１を受信する
400
Ｍ１が望ましくないものであるか？
402
No
完了
Yes
確認されたか？
403
No
Yes
マスタ・シグニチャ・データベースＤ１
Ｍ１を走査する
404
Ｍ１がＤ１内のいずれかのシグニチャ(複数も可)を含むか？
405
No
Yes
Ｍ１がＤ１内のいずれかの原型と一致するか？
406
Yes
No
１つの既存のクラスタがＭ１に十分類似しているか？
410
No
Yes
新しいCaseData要素を計算し、Ｄ１内の一致原型のCaseListに追加する
408
新しいArchetypeData要素を計算し、Ｄ１内のクラスタに追加する
412
Ｍ１用の新しいクラスタ・データを作成する：必要なSigListデータおよびArchetypeDataを計算し、新しいクラスタをＤ１に追加する
416
任意選択でＤ１内のクラスタ用のSigListデータを再計算する
414
ローカル・シグニチャ・データベースを更新する
418

【図５】


メッセージ
M2
メール・コーパス内の
各ｎグラムのオカレンス
の回数をカウントする
500
メール・コーパス内の
ｎグラム度数のデータ
ベースに記憶する
501
メッセージ本文を
抽出する
502
メッセージ本文に不変
変換を適用する
504
ｎグラム度数を使用して、各潜在的
シグニチャの推定の可能性を計算し
最も可能性の低いｍをシグニチャと
して選択する
506
各シグニチャごとに１つまたは
複数のRegionData要素を計算する
508

【図６】


メッセージ
M2
メッセージ本文を
抽出する
602
メッセージ本文に
不変変換を適用する
604
ローカル・シグニチャ
・データベース
D2
変換済みのメッセージ本文を走査して、ローカル・シグニチャ・データベースに含まれるシグニチャ(複数も可)が存在するかどうかを確認する
606
見つかった各シグニチャごとに最長領域一致を見つける
608
HashBlock類似性を
計算する
610
612
一致レベルを決定する
ユーザ
プリファレンス
削除、フラグ付け、通知などの適切なアクションを実行する
614
マスタ・シグニチャ・
データベースを更新する
616